



PHD

Traffic Lights and Smiley Faces

Do children learn mathematics better with affective Open-Learner Modelling tutors?

Girard, Sylvie

Award date:
2011

Awarding institution:
University of Bath

[Link to publication](#)

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

Take down policy

If you consider content within Bath's Research Portal to be in breach of UK law, please contact: openaccess@bath.ac.uk with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

Traffic Lights and Smiley Faces:

Do children learn mathematics better with
affective Open-Learner Modelling tutors?

Sylvie Anne-Sophie Girard

A thesis submitted for the degree of Doctor of Philosophy

University of Bath
Department of Computer Science
February 2011

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with its author. A copy of this thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and they must not copy it or use material from it except as permitted by law or with the consent of the author.

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation.

*The impact of Human Factors and Affect on user's behaviour has to be considered closely
for every bit of technology manipulated by users of all age, but especially children...
It is for us, researchers, to explore the technology's potential to improve children's experiences
when using such products, and to welcome the adventure they lead us in,
when building technologies with and for child users.*

*For every child who thinks doing maths is fun,
and the teachers who try to convince the rest of them...*

THESIS ABSTRACT

This PhD thesis investigates how the use of open-learner modelling (OLM) techniques and the inclusion of affective components in the design of intelligent learning environments can facilitate learning and enhance software usability by increasing children's motivation and engagement in the learning process. The research solely focuses on mathematical applications, given to English and French children aged seven to nine years.

The main contribution of this PhD concerns the study of children's willingness and ability to use affective OLM applications for better learning. The results show that the way children interact with an OLM application depends on its level of openness and student's control over the learning process. Children seem to want to access their learner model components. Such access, as well as the understanding of the learner model content, is facilitated by affective embodied pedagogical agents. The children using an intelligent tutoring system with a negotiated learner model appeared to learn more than children who used an environment with an editable or inspectable learner model, as their learning gain from during each learning session on software revealed to be higher. The use of two different representations of the learner model content - one representing the children's self-beliefs, and the other the system's assessment of knowledge acquisitions – has proven to lead children to be more involved in the representation of what they know by visually comparing their views of how much a specific concept is grasped to the system's assessment, and engaging in a negotiation process when a disagreement was found, which led them to learn more from the sessions on software.

The results and contributions of this thesis should help give evidence of which theories of emotions better apply to children aged seven to eleven working on OLM applications, how children can, want, and effectively use learner model components according to its representation, content, and method of interaction, and therefore help in the design of future affective OLM educational applications for primary school children.

ACKNOWLEDGEMENTS

I would like to acknowledge my supervisor Dr Hilary Johnson, whose guidance and encouragement, infallible support, extensive knowledge on the problem domain, and understanding of the time constraints encountered for the completion of the PhD, made this adventure possible. Thanks for making this whole experience worth it and a true discovery of self.

I would also like to thank the members of my PhD scientific committee, and the members of my PhD viva committee, Drs. Leon Watts and Janet Read, and Pr. Peter Johnson, whose valuable suggestions and comments during the progress reports helped focus this PhD to hopefully make a more coherent and consistent work, taking into account various research domains and expertises.

I would like to extend my thanks to the teachers, evaluators, observers, and the children from all primary schools involved in the various studies performed in this work. Without your help, nothing would have been possible.

I thank the colleagues from different countries and various research domains, met at work, conferences or workshops, for the interesting perspectives and scientific discussions that helped give this research some depth.

Last, but not least, I acknowledge my family and friends who, thanks to their emotional support and care, helped me stay focussed and made this work possible. Mum, dad, Claire, Yannick, Sybille, Gaby, Sabah, Anastasia, and all the others, thanks for being always there, whatever happens.

LIST OF PUBLICATIONS

International journal articles with program comity:

1. Girard, S., and Johnson, H. (2009). « Design affective educational software products: Sorémo, a new method for capturing emotional states ». Journal of Engineering Design, Special Issue on Design and Emotion. 1466-1837, Volume 20, Issue 5, 2009, Pages 493 – 510

International Conference Articles with program comity:

2. Girard, S., and Johnson, H. (2010). Designing Affective Computing Learning Companions with Teachers as Design Partners. 3rd International Workshop on Affective Interaction in Natural Environments (AFFINE'10) at ACM Multimedia 2010, pp49-54

3. Girard, S. and Johnson, H. (2010). What Do Children Favor as Embodied Pedagogical Agents? In. Intelligent Tutoring Systems 2010, V. Aleven, J. Kay, and J. Mostow (Springer Eds.), Part I, LNCS 6094, pp. 307–316, 2010

4. Girard, S. and Johnson, H. (2008). Designing and evaluating affective open-learner modeling tutors. In Proceedings of the 7th International Conference on Interaction Design and Children (Chicago, Illinois, June 11-13, 2008). IDC'08. ACM, New York, NY, pp. 13-16

5. Girard, S. and Johnson, H. (2008). Towards Guidelines for Designing Open-Learner Modeling Tutoring Systems: Mixing Colors and Emotions. In the proceedings of the WEITS workshop, ITS conference, Chicago, IL, pp. 44-54

6. Girard S. (2007), « Towards promoting meta-cognition using emotive interface personas within Open-Learner Modelling Environments », In Young Researchers' Track proceedings of the International Conference on Artificial Intelligence in Education 2007, Los-Angeles, CA. 10 pages

7. Girard S. (2007), « Towards promoting meta-cognition using emotive interface personas within Open-Learner Modelling Environments », (Abstract) In Proceedings of the International Conference on Artificial Intelligence in Education 2007, Los-Angeles, pp. CA, 687-688

8. Girard, S. and Johnson, H. (2007). Including affect in Open-Learner Modelling environments: the impacts of emotive interface personas on learning. In Emotion in HCI: Joint Proceedings of the 2005, 2006, and 2007 International Workshops, 2008, pp. 180-182, Fraunhofer IRB Verlag Stuttgart Ed., ISBN: 978-3-8167-7540-9

9. Fabri, M., C. Peter, S. Girard and H. Agius (2007). An affective application for tackling patient depression. In Emotion in HCI: Joint Proceedings of the 2005, 2006, and 2007 International Workshops, 2008, pp. 199-203, Fraunhofer IRB Verlag Stuttgart Ed., ISBN: 978-3-8167-7540-9

National conferences Articles with program committee:

10. Girard, S. and Johnson, H. 2011. "Designing affective animations with children as design partners using role-playing". In Proceedings of the 23rd international Conference on Association Francophone D'interaction Homme-Machine (Sofia-Antipolis, France, October 24 - 27, 2011). IHM '11. ACM, New York, NY.

11. Girard, S. and Johnson, H. 2009. "Educational software for children: analysis of interaction techniques for direct manipulation". In Proceedings of the 21st international Conference on Association Francophone D'interaction Homme-Machine (Grenoble, France, October 13 - 16, 2009). IHM '09. ACM, New York, NY, 259-262. *Won the Best short paper award*

12. Girard, S. and Johnson, H. (2008). Vers les règles ergonomiques pour les logiciels éducatifs avec une modélisation ouverte à l'apprenant: mélanger couleurs et émotions. *In Actes de la conférence ERGOIA, Biarritz, 15-17 Octobre 2008, ISBN 2-9514772-7-9, pp.177-182*

13. Girard S., Johnson H., (2007). « Opening the Learner-Model to teachers », *In Actes de la conférence EIAH 2007, Lausanne 27-29 Juin 2007, pp. 329-334*

Technical Reports:

14. Girard, S. and Johnson, H. (2007). Opening the Learner-Model to teachers. Technical Report CSBU-2007-20, Department of Computer Science, University of Bath. ISSN 1740-9497

(English translation of the French EIAH'07 conference article)

15. Girard, S. and Johnson, H. (2006). DividingQuest: Using emotive interface personas in educational software. Technical Report CSBU-2006-20, Department of Computer Science, University of Bath. ISSN 1740-9497

16. Girard, S. (2006). DividingQuest: Using emotive interface personas in educational software. Technical Report CSBU-2006-19, MSc. Thesis, Department of Computer Science, University of Bath

TABLE OF CONTENTS

INTRODUCTION	1
RESEARCH AIMS	1
MOTIVATION FOR RESEARCH INTO AFFECTIVE OLM TUTORS	1
SOME ISSUES IN AFFECTIVE OLM TUTORS	3
<i>Design and Evaluation of OLM Tutors</i>	4
<i>Designing, Understanding, and Expressing Emotions</i>	4
<i>Representation of, Understanding of, and Interaction with the Learner Model</i>	4
RESEARCH QUESTIONS TO BE ADDRESSED	5
STRUCTURE OF THE THESIS	6
 CHAPTER 1. LITERATURE REVIEW	 9
1.1 INTRODUCTION	11
1.2 CHILDREN, LEARNING PROCESSES, AND MATHEMATICS	13
1.2.1 <i>Child Development and Learning Theories</i>	14
1.2.2 <i>Understanding, Human Memory, and Motivation to Learn</i>	18
1.2.3 <i>Learning and Motivation</i>	20
1.2.4 <i>Learning and Teaching Mathematics with Technology</i>	23
1.3 EMOTION, CHILDREN, AND LEARNING	25
1.3.1 <i>Child Emotional Development Theories</i>	25
1.3.2 <i>Interplay between Emotions and Learning</i>	29
1.3.3 <i>Capturing young users' Emotional States</i>	38
1.3.4 <i>Emotions for Learning: Discussion</i>	43
1.4 OPEN-LEARNER MODELING ENVIRONMENTS	46
1.4.1 <i>Using Technology to teach Mathematics in Classrooms</i>	46
1.4.2 <i>Intelligent Tutoring Systems</i>	48
1.4.3 <i>Opening the Model to Students: Advantages and Issues</i>	51
1.4.4 <i>Representation of, and Interaction with the model: Using affective EPAs as Interaction Medium.</i>	61
1.5 HCI FOR TECHNOLOGY-ENHANCED LEARNING	67
1.5.1 <i>From Human-Computer Interaction to Child-Computer Interaction</i>	67
1.5.2 <i>Children as Design Partners</i>	70
1.5.3 <i>Evaluating systems with and for children: the Importance of Fun</i>	73
1.6 SUMMARY	75

CHAPTER 2: TWO OLM PLATFORMS: DIVIDINGQUEST AND MULTIPLIOTEST	77
2.1 INTRODUCTION	79
2.2 SOFTWARE ARCHITECTURE	79
2.3 HCI-RELATED ISSUES	82
2.3.1 <i>Designing software for French and English Users</i>	83
2.3.2 <i>Design of the Applications</i>	83
2.3.3 <i>Traffic-Light System and Smiley Faces : Definition of the Metaphors</i>	84
2.3.4 <i>Embodied Pedagogical Agents</i>	85
2.3.5 <i>interaction Style and Usability</i>	87
2.4 AFFECTICE COMPUTING IN OLM APPLICATIONS	89
2.4.1 <i>Sorémo, Evaluating Children's Emotional State</i>	89
2.4.2 <i>Design of the A.M.E.R. model</i>	97
2.5 OLM COMPONENTS DESCRIPTION	107
2.5.1 <i>Pedagogical Component</i>	108
2.5.2 <i>Leaner Model Component</i>	112
2.5.3 <i>User-Interface Component</i>	122
2.5 SUMMARY	130
 CHAPTER 3: EMOTION, CHILDREN, AND LEARNING	 131
3.1 INTRODUCTION	133
3.2 EMBEDDING THE MODEL IN THE DIVIDINGQUEST	134
3.2.1 <i>Techniques used and Rationale</i>	134
3.2.2 <i>Design Procedure</i>	135
3.2.3 <i>Design Process</i>	136
3.2.4 <i>Specifics of the Method's Design to maximize Children's Involvement</i>	140
3.2.5 <i>Design Recommendations</i>	142
3.3 TESTING THE MODEL'S VALIDITY	144
3.3.1 <i>Aim of the Pilot Study and Procedure</i>	144
3.3.2 <i>Study Design</i>	145
3.3.3 <i>Results</i>	146
3.3.4 <i>Conclusion</i>	148
3.4 SUMMARY	149

CHAPTER 4: DESIGNING THE CONTENT OF THE LEARNER MODEL **151**

4.1	INTRODUCTION	153
4.2	WHAT REPRESENTATIONS OF LEARNER MODELS DO CHILDREN FAVOUR AND BEST UNDERSTAND?	153
4.3	LM VIEWS AND REPRESENTATION TOOLS	154
4.3.1	<i>Items of LM content: LM representation tools</i>	154
4.3.2	<i>Organisation of the LM components: LM views</i>	156
4.4	STUDY DESIGN AND PROCEDURE	157
4.4.1	<i>Participants</i>	157
4.4.2	<i>Research questions</i>	157
4.4.3	<i>Procedure</i>	157
4.4.4	<i>Experimental measures and analysis</i>	161
4.5	RESULTS	162
4.5.1	<i>STUDY IV.1: use of visual representation tools</i>	162
4.5.2	<i>STUDY IV.2: structure of the learner model: the LM views</i>	167
4.6	DISCUSSION	171
4.4	SUMMARY	173

CHAPTER 5: OPEN OR CLOSED LEARNER MODEL? **175**

5.1	INTRODUCTION	175
5.2	AIMS OF THE EVALUATION	177
5.3	STUDY DESIGN	178
5.3.1	<i>Research Hypotheses</i>	178
5.3.2	<i>Participants</i>	179
5.3.3	<i>Experimental Design</i>	179
5.3.4	<i>Experimental Measures</i>	179
5.3.5	<i>Procedure</i>	182
5.3.6	<i>Pre- and Post- session briefing of the participants</i>	183
5.3.7	<i>Choice of statistical analysis for research hypotheses</i>	185
5.4	RESULTS	186
5.5	DISCUSSION	193
5.6	SUMMARY	195

CHAPTER 6: AFFECTIVE EMBODIED PEDAGOGICAL AGENTS AS LM-INTERACTION MEDIUM **197**

6.1	INTRODUCTION	199
6.2	AIMS OF THE EVALUATION	200
6.3	STUDY CONSIDERATIONS	201
6.3.1	<i>The affective experimental condition</i>	201
6.3.2	<i>OLM considerations</i>	202
6.4	STUDY DESIGN	203
6.4.1	<i>Research Hypothesis</i>	203
6.4.2	<i>Participants</i>	204
6.4.3	<i>Experimental Design</i>	205
6.4.4	<i>Experimental Measures</i>	205
6.4.5	<i>Procedure</i>	207
6.4.6	<i>Choice of Statistical Analysis for Research Hypothesis</i>	208
6.5	RESULTS	209
6.6	DISCUSSION	212
6.7	SUMMARY	215

CHAPTER 7: INSPECTABLE, EDITABLE, OR NEGOTIATED LEARNER-MODEL? **217**

7.1	INTRODUCTION	219
7.2	AIMS OF THE EVALUATION	221
7.2.1	<i>Research Aims</i>	221
7.2.2	<i>OLM Considerations</i>	222
7.3	STUDY DESIGN	223
7.3.1	<i>Research Hypothesis</i>	223
7.3.2	<i>Participants and Experimental Design</i>	224
7.3.3	<i>Experimental Measures</i>	226
7.3.4	<i>Procedure</i>	228
7.3.5	<i>Choice of Statistical Analysis for Research Hypothesis</i>	229
7.4	RESULTS	230
7.5	DISCUSSION	243
7.6	SUMMARY	247

CHAPTER 8: CONCLUSION AND OUTCOMES	249
8.1 THESIS SUMMARY	249
8.1.1 <i>Experimental Platforms</i>	250
8.1.2 <i>Experimental Results</i>	254
8.2 CONTRIBUTIONS	257
8.2.1 <i>Experimental Platforms</i>	257
8.2.2 <i>Learning in Mathematical Drill-and-Practice OLM applications</i>	258
8.2.3 <i>Application of theories of emotions for children's learning using technology</i>	260
8.2.4 <i>Participatory-design of affective technologies with children as design partners</i>	263
8.2.5 <i>Field evaluations of pedagogical applications in classroom settings</i>	265
8.2.6 <i>Children's use of the leaner model</i>	272
8.3 FUTURE WORK	276
8.3.1 <i>Applications and Feasible Research Directions with the Current Systems</i>	276
8.3.2 <i>Long Term Research Directions</i>	277
 REFERENCES	 279
 APPENDICES	 295

LIST OF FIGURES

FIGURE I.1: THE DIFFERENT STEPS OF THE LEARNING PROCESS, IN (PNS, 2007)	20
FIGURE I.2: AFFECT CIRCUMPLEX (BARRETT & RUSSELL, 1998)	32
FIGURE I.3: FOUR QUADRANTS MODEL OF EMOTION	35
FIGURE I.3: FACS FACIAL READING POINTS	39
FIGURE I.4: SELF-ASSESSMENT MANIKIN: THE THREE PAD DIMENSIONS	41
FIGURE I.5: AFFECTBUTTON AND THE NINE PROTOTYPICAL EXPRESSIONS	42
FIGURE I.6: TRADITIONAL ITS ARCHITECTURE	49
FIGURE I.7: DIFFERENT GRAPHICAL REPRESENTATIONS OF LM	61
FIGURE I.8: VIEWS OF THE LM, IN (MABBOTT & BULL, 2004)	62
FIGURE I.9: USE OF BAYESIAN NETWORKS AND EPAS IN OLM, IN (ZAPATA & GREER, 2004)	65
FIGURE I.10: USER-CENTRED ITERATIVE DESIGN	70
FIGURE I.9: ROLES CHILDREN CAN PLAY IN THE DESIGN PROCESS	71
FIGURE II.1: OLM ARCHITECTURE FOR DIVIDINGQUEST AND MULTIPILOTTEST	80
FIGURE II.2: MVC APPLIED TO OLMs	81
FIGURE II.3: THE THREE STATES OF THE TRAFFIC-LIGHT SYSTEM AND THE SMILEY FACES	84
FIGURE II.4: REPRESENTATION OF SORÉMO EMOTIONAL STATES ACCORDING TO THE AXIS OF PLEASANTNESS AND AROUSAL	91
FIGURE II.5: MEAN LEVELS OF PLEASANTNESS AND AROUSAL	92
FIGURE II.6: THE INSTRUMENT INTERFACE	93
FIGURE II.7: SORÉMO PROCEDURE	94
FIGURE II.8 THE SCALING SYSTEM, SORÉMO	95
FIGURE II.9: EPA'S AFFECTIVE RESPONSE MODEL WITHIN THE OLM SYSTEMS	104
FIGURE II.10: DIVIDINGQUEST ACTIVITY MAP	109
FIGURE II.11: DIVIDINGQUEST LEVELS OF DIFFICULTY FOR ONE ACTIVITY	109
FIGURE II.12: DIVIDINGQUEST LEVELS OF DIFFICULTY FOR ONE ACTIVITY	110
FIGURE II.13: DIVIDINGQUEST AND MULTIPILOTTEST SOFTWARE INTERFACE	122
FIGURE IV.14: THE 3-POINTS SCALE COLOUR-O-SMILEYS SEPARATES LEARNING IN THREE STATES.	124
FIGURE IV.15: EXAMPLE OF "BOXES" IN THE "CONCEPT MAP" VIEW FOR EACH INTERFACE	124
FIGURE II.16: "RELATED CONCEPT MAP VIEW" FOR NLM INTERFACE, ACTIVITY OF LEVEL 1	125
FIGURE II.17: PART OF THE SELF-IDENTIFICATIONS IN THE AILM, U-ELM, AND NLM INTERFACES	126
FIGURE II.18: EDITION OF THE SELF-IDENTIFICATIONS, U-ELM INTERFACE	127
FIGURE II.19: MULTIPLIO FEEDBACK FOR THE AILM INTERFACES IN CASE OF WRONG ANSWER	127

FIGURE II.20: MULTIPIO FEEDBACK FOR THE AILM INTERFACES IN CASE OF A "GOOD TRY" ACTIVITY	128
FIGURE II.21: U-ELM INTERFACE, GIVE YOUR OWN FEEDBACK FOR A QUESTION	128
FIGURE II.22: NEGOTIATION PROCESS, GIVE YOUR OPINION FOR NLM INTERFACE	129
FIGURE II.23: PROGRESS BAR FOR NLM INTERFACE	129
FIGURE III.1: USING FACS TO DESIGN FACES.	138
FIGURE III.2: PEDAGOGICAL SMILEY-O-METER	139
FIGURE IV.1: THE SIX LM TOOLS USED (SKILL-METER, 2D-GRAPH, BOXES, TABLES, TEXT, AND BAYESIAN GRAPH)	154
FIGURE IV.2: RELATED CONCEPTS VIEW OF A LEARNER MODEL USING BAYESIAN GRAPHS	155
FIGURE IV.3: LEARNER MODEL REPRESENTED BY SKILL-METERS UNDER FOUR LM VIEWS.	156
FIGURE IV.4: 'MODEL 1' REPRESENTATION WITH SKILL-METERS.	158
FIGURE IV.5: CHILD'S PREFERENCE OF LM TOOLS, AN EXAMPLE	160
FIGURE IV.6: REPRESENTATION OF CHILDREN'S UNDERSTANDING OF LM VISUALIZATION TOOL ACROSS ALL AGE-GROUPS	163
FIGURE IV.7: REPRESENTATION OF CHILDREN'S UNDERSTANDING OF « SKILL METER » BY AGE GROUP.	164
FIGURE IV.8: REPRESENTATION OF CHILDREN'S UNDERSTANDING OF « BOXES » BY AGE GROUP.	165
FIGURE IV.9: CHILDREN'S PREFERENCE OF VISUAL REPRESENTATION TOOLS	165
FIGURE IV.10: REPRESENTATION OF CHILDREN'S UNDERSTANDING OF LEARNER MODEL THROUGH A LM VIEW.	168
FIGURE IV.11: REPRESENTATION OF CHILDREN'S UNDERSTANDING UNDER THE "LECTURE" VIEW BY VISUAL REPRESENTATION TOOL.	169
FIGURE IV.12: REPRESENTATION OF CHILDREN'S UNDERSTANDING UNDER THE "RELATED CONCEPTS" VIEW BY VISUAL REPRESENTATION TOOL.	169
FIGURE IV.13: CHILDREN'S PREFERENCE OF LM VIEW	169
FIGURE V.1: PROPORTIONAL LEARNING GAIN ACROSS CONDITIONS	186
FIGURE VII.1: EXPERIMENT DESIGN, THE 3 OLM USER-INTERFACES	225
FIGURE VII.3: AVERAGE LEARNING GAIN (%) PER INTERFACE PER SESSION	230
FIGURE VII.4: MEAN OF THE NUMBER OF LM INSPECTION PER INTERFACE PER SESSION	231
FIGURE VII.5: CHILDREN'S CHOICE OF ACTIVITIES BETWEEN THE AILM AND NLM	232
FIGURE VII.6: MEAN OF THE NUMBER OF LM MODIFICATION PER INTERFACE PER SESSION	233
FIGURE VII.7: CHILDREN'S NEGOTIATION OF THE LM PER COHORT	234
FIGURE VII.8: MEAN OF THE NUMBER OF SELF-REPORTS PER INTERFACE PER SESSION	235
FIGURE VII.9: REPARTITION OF THE EMOTIONAL STATES REPRESENTED AS SINGLE UNIQUE, SUCCESSION OF UNIQUE, OR COMPLEX VIEW OF EMOTIONS	236
FIGURE VII.10: GROUPING OF THE SORÉMO EMOTIONS IN TERMS OF AROUSAL AND PLEASANTNESS DURING LEARNING SESSIONS FOR THE AILM INTERFACE	237
FIGURE VII.11: GROUPING OF THE EMOTIONS IN TERMS OF AROUSAL AND PLEASANTNESS DURING LEARNING SESSIONS FOR THE U-ELM INTERFACE	238
FIGURE VII.12: GROUPING OF THE EMOTIONS IN TERMS OF AROUSAL AND PLEASANTNESS DURING LEARNING SESSIONS FOR THE NLM INTERFACE	239
FIGURE VII.13: REPARTITION OF THE EMOTIONS SCALED IN ANTICIPATION OR AS CURRENTLY EXPERIENCED EMOTIONAL STATE	240

LIST OF TABLES

TABLE I.1: PIAGETAN STAGES OF CHILD DEVELOPMENT	14
TABLE I.2: ACUFF AND REIHER'S STAGES OF CHILD DEVELOPMENT	16
TABLE I.3: CONTROL-VALUE THEORY OF ACHIEVEMENT EMOTIONS	30
TABLE II.1: DESCRIPTION OF THE EPA'S DESIGN	86
TABLE II.2: SORÉMO SCALING IN TERMS OF STRENGTH	91
TABLE II.3: SMILI! OLM FRAMEWORK USED TO DESCRIBE THE DIVIDINGQUEST OLM SOFTWARE: THE GENERAL ISSUES.	112
TABLE II.4: ADAPTATION OF THE SMILI! OLM FRAMEWORK DESCRIBING IN MORE DETAILS THE ELEMENTS OF LEARNER MODEL FROM EACH PIECES OF SOFTWARE.	116
TABLE III.1: EMOTIONS ASSIGNED TO EACH DESIGN GROUP	136
TABLE III.2: REPARTITION OF CHILDREN PER INTERFACE	145
TABLE IV.1: DESCRIPTION OF THE MODELS USED IN STUDIES IV.1 AND IV.2	158
TABLE IV.2: MODELS RECOGNIZED TO A "SATISFACTORY" LEVEL OF UNDERSTANDING, ADDRESSED BY REPRESENTATION TOOL USED, AND AGE OF THE PARTICIPANTS.	162
TABLE IV.3: CHILDREN'S PREFERENCES OF LM TOOLS.	166
TABLE IV.4: MODELS RECOGNIZED TO A "SATISFACTORY" LEVEL OF UNDERSTANDING, ADDRESSED BY REPRESENTATION TOOL USED, AND LM VIEW REPRESENTATION.	167
TABLE IV.5: CHILDREN'S RANKINGS OF LM VIEWS PREFERENCES.	170
TABLE V.2: AVERAGE LEARNING GAIN (%) DURING THE SESSION BY CHILDREN'S LEVEL GROUPS	187
TABLE V.3: ANSWERS CONCERNING THE MOTIVATION (NB ANSWER)	187
TABLE V.4: QUESTIONNAIRE ON LM USE, ILM CONDITION.	189
TABLE V.5: QUESTIONNAIRE ON LM NEEDS, CONTROL CONDITION.	190
TABLE V.6: CHILDREN'S LM INSPECTION ACCORDING TO INTERACTION PATTERNS AND LEVEL OF CHILD'S MATHEMATICAL ABILITIES	191
TABLE VI.3: ANSWERS CONCERNING THE MOTIVATION (NB ANSWER)	210
TABLE VI.4: ANSWERS CONCERNING THE EPAs (NB ANSWER)	211
TABLE VII.3: GROUPING OF STUDY PARTICIPANTS BY CONDITION AND AGE.	224
TABLE VII.4: FROM LM INSPECTION TO ACTION	226
TABLE VII.5: CHILDREN'S RESPONSE TO DISCREPANCIES	227

LIST OF ABBREVIATIONS

Abbreviation	Meaning
CCI	Child-Computer Interaction
CHI	Computer-Human Interaction
CLM	Closed-Learner Model(ling)
EPA	Embodied Pedagogical Agent
IDC	Interaction Design for Children
ILM	Inspectable Learner Model(ling)
ITS	Intelligent Tutoring System
LM	Learner Model
NLM	Negotiated Learner Model
OLM	Open-Learner Model(ling)
PD	Participatory-Design

Traffic Lights and Smiley Faces:

Do children learn mathematics better with
affective Open-Learner Modelling tutors?

Introduction

RESEARCH AIMS

The overall goal of this work is to investigate how the use of open-learner modelling (OLM) and the inclusion of affective components in the design of intelligent learning environments can facilitate learning. The impact of such OLM components will also be evaluated in terms of increasing children's motivation to use software, and their use of learner model information in the learning process.

MOTIVATION FOR RESEARCH INTO AFFECTIVE OLM TUTORS

The increasing use of today's technologies in everyday-life activities has led to significant changes in teaching and learning practices as well as the creation of technology-oriented courses in schools such as ICT (e.g. Information, Communication, Technology) (Woolf et al, 2010b). Technological devices are commonly used in the teaching of core curriculum subjects such as mathematics as a medium to facilitate learning (Arroyo et al 2010).

The use of such technologies in classroom settings transforms both teachers' and children's interactions with technology as well as children's learning experiences (Markopoulos et al, 2008): while user-centred design techniques help in the production of applications that respect teaching practices, children's abilities in using technology might influence their ability to perform with software, and change the way tasks are designed. It is therefore important to understand the impact digital technologies have on children's ability to realize pedagogical tasks in order to better support them in their learning process in using such tools. It is also essential to design learning environments that are not only meaningful and useful, but also motivate them to engage more in the learning process.

Research in educational technology has been increasingly interested in gaining some better insights into the development of emotional intelligence, the emotional interactions elicited by using technology to learn, and the interplay between emotions and learning. The inclusion of affect in user-interfaces such as the simulation of emotional expressions by computer agents has been shown to

create an emotional interaction between the user and its agents (Baylor, 2005). Clearly, emotional responses influence the way a person appropriates a product, and its use in the long-term (Holbrook, 1985). However, some issues remain in identifying the effects on learning and motivation using affective computing in learning environments. Currently, there is also a lack of affective methods to help children express their emotions in using products of various kinds, but particularly software products.

Computer-based learning environments seem to be potentially of great value in their capacity for producing personalised learning environments as well as increasing children's motivation in learning (Woolf et al, 2010a). However, a significant number of such existent systems do not suit the teachers' educational needs or technological abilities, due to a lack of teacher's involvement in the design process, and a lack of conformity to existing teaching practices (Van Dam et al, 2005).

Research in educational technology has been interested in the creation of environments that enable the most able children to be stretched academically, whilst providing extra-support at the basics for those who require it. In order to reduce the gap between school practices and the interactive educational activities proposed by tutoring systems, there is a need for designers to perform a more extensive investigation of how to mirror/adapt such practices by taking advantage of the possibilities offered by technology, as well as its related constraints (Markopoulos et al, 2008).

The use of user-adaptive systems, designed in close collaboration with teachers and children, have been considered to enable individual learning through user-modelling techniques (Bull & Kay, 2010). Part of the user-modelling community is concerned with the creation of Open Learner Modelling (OLM) systems, or systems that enable access and/or inspection of its students user-model to the students as well as their teachers, parents or peers. These modelling techniques have the potential to create systems that are useful for children to learn a particular subject, by providing children with some reflective tools to think about their learning process (Zapata-Riviera, 2003). They also could facilitate the creation of highly usable for both children and teachers and able to be integrated into traditional teaching strategies.

On the one hand, giving access for a learner to his/her model or his/her peer's user model has been found to improve the learning skills and performance of the learner for teenagers and adults (Kay & Li, 2006; Kerly & Bull, 2006; Zapata-Riviera, 2003). On the other hand, the creation of Open-Learner Modelling systems for children and their impacts on their learning have yet to be defined. Indeed, the inspection/modification of the learner model (LM) content offered to the child-users can help learners gain a better understanding and conceptualization of their achievements as regards to the learning goals present in the curriculum, and even play a role in the way they plan/orchestrate their learning sessions (Bull & Kay, 2010; Kay, 2001). However, several issues have been outlined in terms

of understanding and accuracy of the learner model content, as well as student's willingness to use such information (Barnard & Sandberg, 1996; Bull & Kay, 2010; Kay, 1995).

Investigating ways to improve the design and creation of OLM tutoring systems for children through the use of affective computing, and student modelling therefore seems of great research value. This thesis will concentrate on exploring how children understand and (want to) use the learner model in OLM Intelligent Tutoring Systems (ITS). It will investigate the potential of using affective feedback and different ways to present and interact with the learner model information. This is accomplished through the design of an ITS to increase learning gain, and children's motivation to use technology to learn.

SOME ISSUES IN AFFECTIVE OLM TUTORS

Research in Affective OLM tutors for use by children in classroom settings raise a number of issues, among which are:

DESIGN AND EVALUATION OF OLM TUTORS

Tutors will be designed to be used in classrooms, as a support tool for teachers to give a more personalized learning session to the children adapted to their level and learning pace. Some issues are raised as to the design of such tools as easily deployable in schools, and with a human-computer interface (HCI) suitable for the children in terms of usability and pedagogical assimilation with the teachers' and learners' practices. How can you involve teachers in the design of such an application to guarantee an appropriate pedagogical content? What place do children hold in the design partnership of the application: what control should they have over the learning process, how can the design choices be made between children's will and learning considerations when they differ?

DESIGNING, UNDERSTANDING, AND EXPRESSING EMOTIONS

The OLM systems under consideration also integrate an emotional component. Research in technology-enhanced learning has been increasingly interested in the impact of the inclusion of affect in the design of pedagogical tutors in terms of emotional feedback from the system and users' responses. However, some issues remain as to the simulation of emotions and the impact of an affective dialogue in terms of learning. Some studies have shown the potential of using pedagogical agents as an interaction medium to give emotional feedback and produce emotional responses (Woolf

et al, 2010a; Zapata-Riviera, 2003), but its impacts on learning when used by children are still unclear. Research in child developmental psychology defined several theories as to how children recognize, understand, manipulate, and express emotions according to their emotional development. The interplay between emotions and learning has also been investigated in the past. However, results as to the impact of using psychological and computational theories to produce emotional feedback in educational technologies, and its impact on user's learning processes and motivation, are yet to be systematically investigated.

REPRESENTATION OF, UNDERSTANDING OF, AND INTERACTION WITH THE USER MODEL

Research in open-learner modelling often involves adult users. When considering child users, some issues are raised as to the ways in which the model can be represented for the children to understand what it contains and represents. Once this model is understood, the question of the user's interaction with such a model remains: Do children want to give their input? What are the benefits at interacting with the learner model to the child-users? Will this change how they learn and conceive learning with technology? Will it give them more insight into their current and future knowledge acquisition, or just impact their short-term goal achievements? How much input is suitable to maximise the learning experience? Can children make use of the OLM by having access to a conceptualization of their level from the system's view only, or do they need help to compare their own mental model representation of skills and achievements to the one presented by the system? Studies have shown that the way the content of the learner model is presented to users changes their willingness to interact with the model, and their understanding of its content. Is there a particular learner model representation that would help the understanding of child-users of their learner model?

The use of embodied-pedagogical agents (EPAs) in educational applications was also reported as increasing the level of user's enjoyment, as well as understanding of the user-interface and/or pedagogical components in learning with technology in intelligent tutoring systems. In this thesis, the focus of the research will be interested in building affective OLM tutors. Embodied Pedagogical Agents expressing affective states according to the level of the child, his/her achievements, and the content of his/her learner model will be used as an interaction medium with the learner model component. It aims at facilitating children's understanding of the model content better, lead to better learning, and increase their motivation to learn using technology.

RESEARCH QUESTIONS TO BE ADDRESSED:

From the study of current research on all aspects of the problem domain, a number of research questions have been drawn, to be answered in this PhD thesis:

RQ1: What theories of emotion can be followed to produce a model of EPA's affective response in OLM software suitable for mathematical drill-and-practice applications?

RQ2: How can user-centred and participatory design techniques be used with children and their teachers to create affective components to be integrated in tutoring systems for children?

RQ3: What is an appropriate representation of a learner-model (LM) for a child to interact with?

RQ4: How can user-centred and participatory design techniques be used on children and their teachers to represent the Learning model components of OLM tutoring systems for children?

RQ5: Does opening the content of the learner model to the child-user facilitate learning and motivation?

RQ6: How can the use of pedagogical agents or the inclusion of affect in the design of OLM tutoring systems help children interact with their LM, facilitate learning and motivation?

RQ7: How can children interact with their learner-model?

RQ8: What level of inspection and modification of the learner model content by the child-users is suitable for the child to learn efficiently?

RQ9: What kind of learner model inspection/modification by users is most suitable to help children make use of the learner model to learn more efficiently and gain more reflection on the pedagogical activity: full editing (editable learner model), no editing (inspectable learner model), or building the model with a user-system negotiation process (negotiated learner model)?

STRUCTURE OF THE THESIS

This PhD thesis is structured in five parts: First, the reader is taken through a discussion of the rationale for the intended research in chapter 1, and the definition of a number of research questions to be addressed in this thesis. These questions necessitated the development of educational software (chapter 2), methods for measuring affect (chapter 3), and a definition of learner model representations for children use (chapter 4). Once finalised, the remainder of the thesis used these methods and tools as a vehicle for undertaking empirical studies to answer the research questions (chapters 5 through 7).

Chapter 1 motivates the research reported here, by providing a literature review of the various research domains, and highlighting the current research issues concerning the development of OLM educational technologies that better help children learn mathematics. It includes a literature review of theories of children's learning and human processes, learning applied to mathematics using technologies; theories on children's emotional development and interplay between learning and affect, as well as its applications in affective computing; the issues of (open-) learner modelling in Intelligent Tutoring Systems in terms of representation, access to, and interaction with the learner model; and the design and evaluation of educational applications for and with children.

Chapter 2 presents all aspects of the two OLM software applications, *DividingQuest* and *Multiplotest*, created to enable investigations aimed at investigating the previously stated research questions. After describing the software architecture, and some considerations of HCI-related issues, each component of the OLM is described and associated with the research presented in the following chapters.

Chapter 3 provides the design and preliminary tests of a model of affective responses from EPAs to user-interaction, created in participatory-design with teachers and following the current theories on emotions for learning. It also includes a description of the participatory-design method that led to the creation of the affective behaviour of the *DividingQuest*'s EPAs, by using children as design partners. The model of emotion is then assessed in terms of its impact on learning in chapter 6, and embedded in every experimental condition of *Multiplotest* in the chapter 7 experimentation.

Chapter 4 is also concerned with the issues regarding the design of OLM educational applications, and studies the aspects of learner model visual representation. It presents a study testing the usability of different ways to represent and display learner model, for children aged seven to eleven to understand the model and use it better. The results of the study are then applied in the construction of the LM representation in the two OLM built, which are illustrated with a description of the transformation of *Multiplotest* to suit the investigations of chapter 7.

This PhD is focussed on the investigation of children's ability and willingness to use an affective OLM tutoring system to learn and practice mathematical subjects, and its impacts on children's learning environments. The experimental part of the thesis is divided into three different experimental studies, described in chapters 5 through 7. In chapter 5, a first research question is addressed, investigating the benefits of opening the learner model to children in terms of learning gain and motivation, as opposed to a tutoring system where the LM content is hidden (neither inspectable nor editable).

Chapter 6 is then concerned with the benefits of using affective embodied pedagogical agents as an interaction medium with the learner model, in such systems. The affective model of EPA's affective responses to child-software interaction, described in chapter 3, is tested against EPAs with lack of affective response, and an interface without any EPA or affective feedback. The results of an experimental study, comparing three versions of the first educational software developed for this PhD's work, the *DividingQuest*, which differ in the interaction medium chosen between the user and the LM, is presented: In the first user interface, the interaction with the model is performed through the use of buttons, skill-meters and box LM representations. In the second, an EPA that portrays a constant positive emotional state is in charge of the interaction, while in the last interface the same EPA portrays different emotional states corresponding to the model previously defined. The study evaluates the benefits in terms of children's motivation, learning gain, and use/understanding of the learner model of the affective EPA interaction medium with the LM.

Chapter 7 includes this thesis's main investigation, in the form of a longitudinal study undertaken with children aged seven to eleven years. It investigates three types of learner model openness, in an OLM where the LM is divided between the system's expertise on children's characteristics, and children's self-identifications of their level:

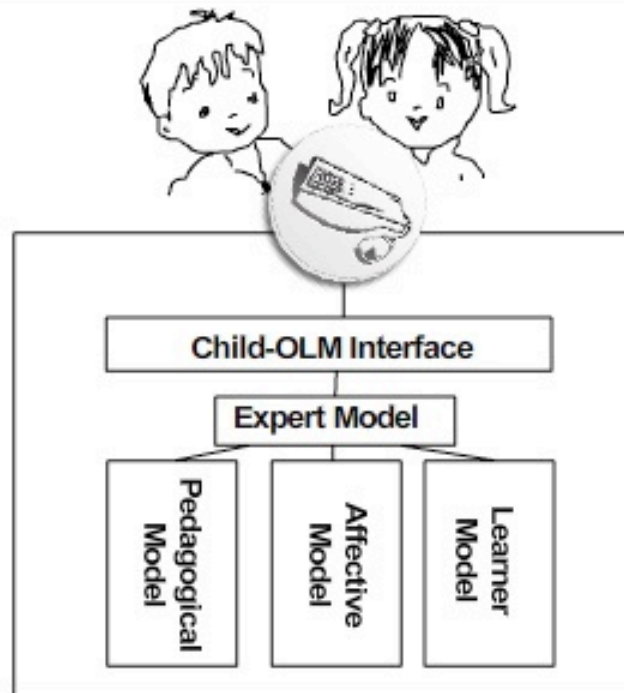
- u-ELM (child's view, user-editable): children's beliefs are inspectable and editable, system LM hidden.
- AILM (system's view, inspectable): only the system LM is inspectable, no edition possible
- NLM (system's and child's views, Negotiated-LM): the model is built each step of the way through negotiation between the two views.

Finally, a conclusion chapter summarises the main aspects of the research, points out the main contributions of this thesis, outlines the limitations of the results, and sketches out directions for future work.

Chapter 1. Literature Review

Children, Learning process, and Mathematics

Open-Learner Modelling Environments



Emotion, Children, and Learning

HCI for Technology-Enhanced Learning

1.1 INTRODUCTION

The use of computers is becoming more frequent in every field of human activity and impacts people at every age and in the majority of occupations. In light of this information revolution, it is essential that children have positive experiences with computers and technology to facilitate their interaction with technology in future life (Wyeth & Wyeth, 2001). Whilst Druin and Inkpen (2001) state that the use of computers can help support creativity and provide social experiences, it is important to study the impact of using computer-based applications on children's learning processes.

Mathematics is a core subject of primary-school curriculum, and has benefited from special interest in adapting learning practices to technological use. However, while some research showed the potential of using intelligent tutoring systems for the learning of core-subjects such as mathematics (Aleven et al, 2001; Arroyo et al, 2010; Berkovitz, 1994), not every child gains from such software use, for various reasons: some interactive components appeal, are usable, or understood by a small number of final users (Markopoulos et al, 2008); software is not properly introduced to teachers; is too far from the existing learning and teaching practices (Van Dam et al, 2005); or is not adapted to the material readily available in classrooms.

Research into product design has clearly demonstrated that products can elicit strong emotional responses, which influence the way a person appropriates the product, and its use in the long-term (Holbrook, 1985). This evaluation of emotions can be seen to affect the way in which children interact with educational software products, with positive emotional responses possibly increasing motivation to both interact and learn (Heraz & Frasson, 2008; Woolf et al, 2010b). The use of pedagogical agents, embedded within pedagogical software of various kinds, is known to make the experience more engaging for users (André et al, 1998; Lester et al, 1997). Adding an affective component to such agents has been shown to create an emotional interaction between the user and its agent (Brave et al, 2005; Woolf et al, 2010a). Therefore, investigating the means of integrating affect into pedagogical agents in the design of educational software might help children be more motivated to use software, and understand better the information provided by software on their learning process.

Intelligent tutoring systems (ITS) are computer-based learning environments structured around four main components: the domain knowledge (also called expert system), the task environment, the student model (or learner model), and the pedagogical module. The learner model of an Intelligent Tutoring System is often defined as “a student-instantiated version of the competency model. That is, values in the student model express the assessor's current beliefs about a student's level on variables with the expert model” (Shute & Zapata-Riviera, 2010). Open-Learner Modelling environments are intelligent tutoring systems applications where the learner model is partly accessible for inspection and manipulation by different actors of the learning process (the learner, teachers, parents, peers, etc...).

Opening the information of the learner model to the users for inspection/modification has potential to increase a user's motivation to learn, and may facilitate learning (Zapata-Rivera, 2003; Kerly & Bull, 2006). However, several issues have been outlined in terms of understanding and accuracy of the model content, as well as willingness to use such information (Barnard & Sandberg, 1996; Bull et al, 2009; Kay, 1995) and can even be detrimental for learning to some child-users (Bull et al, 2005). Research aiming at producing systems that fully use the potential of open learner models in facilitating the understanding of this model and its good use within the learning process are therefore needed.

Section 1.2 will review child learning processes, investigating the theories of child developments, the components of learning, the link between learning and motivation, and the value of using computers as a medium of teaching for the apprenticeship of mathematics. Section 1.3 will then explore research concerning the design and impact of the use of Intelligent Tutoring Systems paying particular attention to the impacts of opening the learner model to users on the learning process, especially when pedagogical agents are used as the child-computer interaction medium. In section 1.4, aspects relating to affect are presented, including theories of children's emotional development, the link between affect and learning, and its application in terms of affective computing. Finally, the last section will present the field of Child-Computer Interaction along with its design and evaluation techniques for the design of children's interactive products.

1.2 CHILDREN AND LEARNING PROCESSES

Effective teaching lies in the ability of the teacher to set up a learning experience that makes the children willing to learn and help them grasp the concepts of the curriculum. In order for effective teaching to take place, each pupil must be engaged in the activity of learning. For this to happen, there is a need to understand the cognitive development of the learning process, enabling teachers and educators to tailor the experience/technology to the specific abilities at different developmental phases.

In this section, existing child development theories and their implications for the learning process are reviewed. The processes leading to the creation of some form of understanding, considered as key factor for learning to be successful, are then detailed. In section 1.2.3, the practices of learning mathematics in primary school will be outlined, as well as their needs in terms of learning process and understanding of concepts. Finally, closer attention will be given to motivational factors and their link to learning.

1.2.1 CHILD DEVELOPMENT AND LEARNING THEORIES

When considering how children learn, it is common to identify the group of children addressed by age. This is the grounding of child development theories, which state that children change with age according to five features of development; namely physical, social, emotional, intellectual and language (Sylva & Lunt, 1982). Early research in developmental psychology in children mainly investigated their limitations, assuming that children were restricted in their abilities compared with adults with respect to what they could comprehend and how they reasoned about situations.

Jean Piaget's theory of cognitive development (Piaget, 1970) classifies the key changes in intellectual and language development by providing a definition of a child in terms of age and cognitive skills. He claims that children 'naturally' pass through a sequence of developmental stages through which they attain conceptual knowledge by handling progressively more complex concepts in more complex ways.

Table I.1 illustrates the stages of child development as defined by Piaget:

Table I.1: Piagetan stages of child development

Stage	Ages (yrs)	Key points for interactive product design
Sensorimotor	0-2	Children rely on what their senses perceive – they will be unable to interact with products. Develops an initial egocentric viewpoint that does not enable them to distinguish between themselves and the rest of the world.
Preconceptual thought	2-4	Brief attention span, can only hold onto thing in their memory at a time. Unable to read, they can understand simple instructions. Towards the end of the period, can build conceptual knowledge of mathematical concepts.
Intuitive thought	4-7	Children can use symbols and words. Children can distinguish reality from fantasy. In the latter part they can take into account the viewpoint of others. Will solve problems by empirical testing rather than logical reasoning. Dominated by their perceptions, are likely to be misled by what they see.
Concrete operations	7-11	Children can classify things; understand the notion of reversibility and conservation. Can think logically but not abstractly. The operations of "classification, ordering, construction of the idea of number, spatial and temporal operations" (Piaget 1970) are part of the children skills at that stage.
Formal operations	From 11 yrs	Thinking about ideas, they can consider various solutions without having to act them all out – can deal with hypothetical situations. It allows them to consider many different factors impacting on a given situation using combinational understanding.

One of the main criticisms addressed to Piaget's theory is the rigidity of the stages, claiming that "children in the right context are able to develop certain understanding at a much earlier stage than he suggests" (Kyriacou, 1997), and that Piaget underplays the role of social factors in pupil's cognitive development.

On the one hand, this classification gives a useful design framework, at the start of the design process: by taking into account the general trends of children this age, the research can be focussed on parts of the curriculum and teaching practices that would most likely be within the framework. This should facilitate the building of an experimentation setting that includes learning content and objectives children are able to understand and perform. On the other hand, Piaget's work (1970, 1977) only focuses on the child's cognitive development as regards to his/her own skills. It does not take into account the changes in learning and understanding related to the child's social environment and the learning strategies he has been subjected to that fostered his motivation to learn (Brave et al, 2004).

Vygotsky's theory of teaching and learning (Vygotsky, 1987) is focussed on the social aspects of learning, rejecting the idea that intelligence is fixed. Researching the impact of the aid of an instructor on the child learning process, he defined the concept of '*Zone of Proximal Development*'. The research behind this concept is to analyze the potential of a child to rise above what they can achieve intellectually through collaboration with a teacher, making a transition from the current skills of a child to what the child could achieve by imitating the instructor and finally do by him/herself. He concludes that children perform and develop better with help, and that tasks given to them should rather be assessing their development than their acquired learning.

Bruce (2001) defined a social learning theory based on her analysis of children's representations including play, which helped identify a number of schemas. A schema represents a pattern of repeated actions, and clusters of schema develop into later concepts. Her theory, focussed on creativity, play, and first-hand experiences, evolves around the notion of "free-flow" play: children's play is at its richest when they are able to "wallow" in it. The theory is that by involving children in a form of 'free-flow' play, they become more interested in the learning activities, and learning gains as well as motivation are increased (Bruce, 2001). She identified twelve characteristics, which mark "free-flow" from other kind of play, including: first hand experience, developing rules and props, freely chosen activity, rehearsing recent learning or celebrating learning, imagining the future, pretending, deep involvement and "personal play agenda", and co-ordinated ideas and feelings.

The goal of this PhD work as regards to children's learning development theories is not to introduce a new theory for the age group and nationality of the children population targeted. It is however to investigate how the current theories can be applied in the design of Open-Learner Modelling tutors to facilitate learning.

Acuff and Reiher (1997) have applied Piaget's stages to describing the effect they can have on the design features for interactive systems. Their distinct groups defines the stages of child development as changing interests, humour, characters, contexts and settings, as illustrated in table I.2.

Table I.2: Accuff and Reiher's stages of child development

Stage	Ages (yrs)	Description
Dependency / Exploratory birth	0 - 2	This is a stage of learning, exploration and discovery. Children enjoy repetitive sensori-motor actions and they cannot yet play together. Products for children up to two should be based on simple concepts, give a feeling of safety and stimulate learning e.g. an activity centre, which has the form of a tablet, with pictures, buttons and sliders.
Emerging Autonomy /	3 – 7	In this stage children enjoy fantasy and magic. Children of this age group are fairly self-centred, doing a lot of parallel play. They have a need for simulation, love and safety, though they are developing a greater need for autonomy. Therefore products for children within this group are often within the context of a fantasy world e.g. children have to search for items that enable them to reach a final goal.
Rule / Role	8 – 12	Interests of children in this age group shift gradually from fantasy to reality. They play in pairs and groups and become more interested in competition. Examples of products are laptops or handheld computing devices. These devices look more adult and serious-like than those designed for younger ages. The use of language can be more complex and abstract.
Early and late adolescence	13 +	In this phase children develop their abstract thinking and logical skills. Activities become both more <i>goal</i> and more <i>socially</i> oriented, such as sports, school-clubs and social activities with friends. Adolescents can handle abstract problems and complexity. They relate most strongly with more realistic characters and prefer realistic settings to a fantasy world. Vocabulary size develops to over 40,000 words on adolescence. Products designers for this age group are very similar to those designed for adults. They have to relate to activities that appeal to this age group such as sports and social activities. The projects need to have a realistic look and contribute to the user's self-image. For example, a wide range of mobile phone covers are available with pictures on them, which enable the users to tailor them to his/her intended image.

Children aged seven to eleven are considered as part of Piaget's 'concrete operational state' and Accuff & Reiher's 'rule/role' stage (1977). They represent a stage where children are more autonomous and like to be more in control of their activities. They are less in need of game-like features for learning that include fantasy worlds, and are able to reason logically. Their interest in more adult-like devices for learning should also facilitate their interest in educational software that give them food for thought, and enable them to have some control over their learning choices, therefore giving them more responsibilities, like adults do.

The use of an Open-Learner Modelling tutor could help children engage more in the learning activity and gain more insight of their own level by following the rules of ‘free-flow’ play. The application of such rules to using OLM educational software could facilitate learning for children, by:

- Experiencing the notions first-hand by choosing the activities in the drill-and-practice tutor, according to their current level, and what they are currently in the mood to learn;
- Being in control of their own representation of knowledge acquisition through the learner model information provided;
- Using the pedagogical agents acting as interaction medium with the learner model to form “play-props” and play-pretend by immersing themselves in the environment;
- Using the information from the learner model to show their skills and competencies, and have a better understanding and mental representation of their current skills and learning achievements.

Could the use of a tutor help children from the concrete operational stage understand and learn better? By presenting abstract concepts in multiple ways, could the use of computer-based learning environments help them integrate new concepts and bridge the gap between the mathematical abstraction and the real-world problem application? Could it also help children on their journey to understanding, maximizing the learning process in the zone of proximal development?

Could an “intelligent tutor” containing an open-learner model use the expert model and its understanding of the child’s knowledge to help him/her grasp new pieces of knowledge? Could this help children in their understanding of the notions, and focus on the misconceptions detected during software interaction?

Could elements designed to express affective cues, strategically placed within software components, help children in the visualization of the zone of proximal development and their achievements within this zone? By making software features more ‘friendly’ and transformed into ‘play-props’ objects, could the use of affect into software design help children interact with help software features and OLM information to a greater extent? How would this increase of child-software interaction in terms of content of the system’s model impact children’s learning and motivation to use OLM software?

1.2.2 UNDERSTANDING, HUMAN MEMORY AND MOTIVATION TO LEARN

For a child to be able to learn a concept, it is essential that s/he understands the information, interprets and compares the information to what s/he already knows, and finally integrates the purpose and procedures of the learning content. The following components of children's understanding process are explored in this section: understanding linked to human memory capacity and structure, children's experiences when understanding and the impact of familiarity to facilitate understanding, and how language skills impact understanding.

Norman (1998) explored how human memory functions to discover the best methods to support remembering and understanding. He defined three types of memory: memory for arbitrary things, memory for meaningful relationships and memory through explanations. Children use memory for arbitrary things, which consists of remembering without a deep understanding of the reasons for doing something in their learning of the alphabet with a song for example. When considering memory for meaningful relationships, children aged seven to eleven will be more likely to remember events or concepts when interacting with people or objects they are familiar with. Children's memory of a concept they understood increase when some kind of explanation is involved (Luckin & Hammerton, 2002). However, Richmond (1970) shows that children better memorize the understanding of a learning concept when they experience the concept within familiar settings, and with a close link to familiar words, events, and skills.

It is argued that abstract associations or sequences are difficult to remember without constructing some kind of mental structure of the concept (Norman, 1998). This is illustrated by the learning of the alphabet with a tune with rhythm and rhyme providing an appropriate structure, making it easier to remember. Children of age seven through eleven will already possess some understanding of mathematical concepts but will be faced with learning new abstract ones, not related to their previous experiences. It is therefore essential for learning that a sensible and meaningful structure exists to help children draw relationships between the items to remember and existing knowledge.

According to Richmond (1970), all new experiences must be related to experiences the child already understands. This theory that learning needs to be acquired by experience can be traced back to Confucius (450 BC.) who stated, *"Tell me, and I will forget. Show me, and I may remember. Involve me, and I will understand."* Children have to feel involve into the learning process, to be able to acquire links or connections to real world experiences and have them at their disposal to apply in different contexts (Wehmeyer et al, 2000). Wehmeyer et al (2000) claims that the success of the understanding performance "depends on the familiarity of the element of the problem with

associations already in the mind”, stating that the application of concepts in different contexts is fundamental for learning.

Some findings from the studies of Donaldson (1978) showed that whilst children do not always interpret language as adults do, they are more likely to interpret the meaning of what they are told if they understand the context of the situation being discussed. Vygotsky (1987) and Liebeck (1984) support this theory by emphasizing the need for children to learn by having “authentic situations in which they must solve dilemmas” (Vygotsky, 1987). They advocate that learning metaphors and authentic examples can help to form the required connections between abstract concepts and the real world, by giving them a more “concrete focus” (Donaldson, 1978).

The use of OLM software, by presenting knowledge acquisition under different forms, could highlight the notions to learn and children’s misconceptions. Could this help children interpret the information given in software? Could the presentation of different views of knowledge and knowledge structure help children in the understanding of how such concepts are linked together, and the hierarchy of skills to acquire as favoured by teachers from their teaching practices?

The use of emotions in the characters, whose role it is to guide the user in his learning experience using the software, might help them consider the helper as a character alive. How could this impact the child-helper interaction, and thus children’s ability to connect concepts better with the real world? Could the use of such characters reinforce children’s memory of meaningful relationships by linking together better the concepts manipulated in the explanations given by the helping characters?

The United Kingdom National Numeracy Strategy (DFEE 1999) promotes the expression of learning in real-world contexts for the knowledge to be assimilated efficiently, and the teaching methods to be better understood. Could the use of educational technology help children in the understanding of mathematical concepts and their learning achievements by integrating metaphors of learning, and helping children link the activities to classroom teaching practices?

1.2.3 LEARNING AND TEACHING MATHEMATICS AT SCHOOL

Teaching is not just about textbooks, children must play an active part in learning. As Economides & Elissaver state (2000), “the failure of so many instructional programs has been the result of an emphasis solely on content, with little regard for principles of instructional design to produce effective, efficient and appealing instruction”. Learning is an active, constructive, cumulative and self-regulated process in which the learner plays a critical role. The UK Primary Numeracy Strategy(PNS, 2007) categorises learning in different steps as a repetitive process: teach, practice, apply, and review and assess, as illustrated in Figure I.1:



Figure I.1: the different steps of the learning process, in (PNS, 2007)

When learning occurs in classroom settings, the teacher aims to define the limits of understanding of the learner, in order to successfully pass on knowledge (Gallimore & Tharpe, 1990). When developing any tool or method to help children learn, a thorough analysis of the concepts to learn, and the pedagogical means to achieve learning, has to take place, from education theories to everyday teaching practices.

In Williams et al (2006)’s study on the relationship between teacher’s beliefs about their classroom roles and its influence on how they negotiate classroom contexts and respond to classroom emotional events, there seemed to be connections between what the teachers believed their roles should be in the classroom, how they structured the activities, and the relationships they sought to develop. Example: creating environment where all students should achieve by putting on her ‘cheerleader’ hat to motivate the students to learn using positive emotions to avoid frustration from students, which brought hope (Harris, 1989).

There are many approaches to teaching children, each of which aims to provoke different responses and nurture different modes of thinking within children. Fisher (1995) identified some tools to be used by the teacher as teaching strategies, namely questioning children, creativity, social interaction, communication through group or pair work, peer learning and coaching. In the case of

technology, an emotional link could be drawn from the child's interaction with an emotive pedagogical agent, not considering the helper as system functionality but as a peer or a coach for the learning session.

When considering the apprenticeship of mathematics, specific activities are included in the learning process. Mathematics investigates the manipulation of numbers, algebra, measures, shapes and the handling of data. It is typically associated with problem solving but also involves a certain amount of knowledge retention. The student is required to remember facts, learn and use newly acquired skills, understand conceptual structures and finally apply problem solving strategies and attitudes.

Pupils from the concrete operational stage (seven to eleven yrs) should be more than fluent in the mathematical principles and will be encouraged to apply their knowledge when answering more abstract problems. The UK Primary Numeracy Strategy (PNS, 2007) advocates that children master the basics of multiplications by year 4, and the basics of divisions by year 5. By that stage they should be able to multiply and divide any combination of units, tens, and thousands. The objective of year 6 in UK is the application of the notions of multiplication and division to concrete situations in problem solving activities, and more sophisticated concepts such as the addition, subtraction, and simplification of fractions, and numbers with a decimal point. In year 6 (nine to eleven yrs), children from primary schools based in UK study the mechanisms of divisions under all aspects (dividing short and long numbers, negative numbers, aspects of divisibility and scientific notation, as well as problem solving). Similarly to the DFEE strategy, children from French primary schools are focused on support, consolidation, and extension of knowledge. At the CE1 level (seven to nine yrs), a new mathematical operation is introduced, multiplications. The principles around the concept are studied in depth, and then rehearsed and extended the year after, at CE2 level (nine to eleven yrs). Could the building and use of a tutor that enables the discovery and practice of one specific mathematical concept (here either divisions or multiplications) help children in the apprenticeship of said notion and visualize all aspects of the concept better?

What place should be given to encourage children to reflect on their knowledge acquisition: Reflecting upon the process of an action is useful, and can lead to a better understanding of the concept. However, a child that concentrate too much on the method and how a multiplication works can encounter issues in applying multiplications to concrete problems by losing the focus on what is asked of him/her. In this PhD, the notions to be acquired in software will be restricted to basic blocs of knowledge, in the concepts of multiplication and divisions. In that context, how much reflection from children on their knowledge, and the process learning to the mathematical concepts is needed, and useful in pedagogical applications?

The DFEE (1999) defines 'high-quality' learning as oral, interactive and lively. Its achievement is not only an adoption of simplistic formula of drill or children learning from books by themselves. It invokes exposing children to appropriate material and concepts at pertinent times for their educational development to follow a suitable path. Teaching is therefore about support, consolidation and extension of knowledge (Hammond, 2001). Learning from books in classroom enables teachers to give children such appropriate materials. However, children do not learn at the same pace, and when confronted with over twenty or thirty students at the same time, the choice of activities to perform becomes an issue. By allowing a personalized learning environment, could adaptive educational software help children by providing different sets of activities, according to their learning pace? The practice of learning concepts using an intelligent tutoring system could help knowledge consolidation and even extend it by complementing the classroom teaching strategies. The use of an affective helper to point out the concepts not yet acquired could help in supporting and consolidating learning by giving specific advice and individualized learning goals on a particular topic.

Questioning is an important part of all mathematics lessons and can be used either to challenge a child to recall information, or when building on previous knowledge to introduce new ideas. Whereas the former helps in building confidence and maintaining a consistent level of child participation in the lesson, the latter requires the child to think about new ideas and possibly apply their knowledge abstractly. A pedagogical agent embedded in OLM software could provide visualization of, and some level of involvement (inspection/edition) in the representation of children's conquered skills, new material to learn, children's current misconceptions, and current/future pedagogical goals to achieve. Could an emotional link between the agent and the child, provided by a model of affective feedback from the agent to user-software interaction, help children in the understanding of the mathematical concepts within software, as well as the concept's organizational structure, therefore creating questioning? Could such interaction with software help users keep focussed on the pedagogical tasks offered by software, and increase their level of motivation?

1.2.4 LEARNING AND MOTIVATION

A prime concern to all teachers is the motivation of the learner. Newton (2001) argued that “a press for understanding is likely to be accompanied by (...) some provision that motivates the children to engage in these processes.” Yet, motivating a child to learn is still, and always will be, an issue.

Copeland (1979) explains that motivation does not only come from the child, but also from his/her interaction with the learning environment. He states that motivation is achieved if the experiences described in the learning session are closely matched to the child’s ability to respond to them. A manipulation of the environment variables to expand upon existing knowledge and set of experiences might therefore result in an increase of motivation.

One way of motivating children to learn is the use of some kind of reward in case of success. In their experiment, Anderson et al (1995) demonstrated that verbal praise made the children from the concrete operational stage spend more time practising the activities and therefore benefited the tutor and the learning process.

According to Keller (1984), there are four factors necessary to achieve optimal motivation: attention, relevance, confidence, and satisfaction. Keller (1983, 1987) presented a series of steps towards improving learner’s motivation, using attention modelling as a first step in detecting motivation. Five strategies were defined, aimed at improving or sustaining the learner’s attention.

To increase attention,

- Use novel, incongruous and paradoxical events. Attention is aroused when there is an abrupt change in the status quo;
- Use anecdotes and other devices for injecting personal, emotional element into otherwise purely intellectual or procedural material;
- Give people the opportunity to learn more about things they already know about or believe in, but also give a moderate dose of the unfamiliar and unexpected;
- Use analogies to make the strange familiar and the familiar strange;
- Guide students into a process of question generation and inquiry: action, as well as attention, have to play an active role in the process.

Could the use of a computer-based tutor help achieve a good level of motivation in doing their learning activities? As the interactivity of the software could draw the child's attention to particular aspects of the concepts learned, could a child also explore unfamiliar problems and show an evolution of the concepts according to difficult contexts? As previously stated, children need to relate to previous experiences to facilitate understanding. Some metaphors of learning are widely used in classroom settings to represent knowledge acquisition, and therefore became familiar to children. Could the digital modelling of metaphors of learning used in classroom based on coloured or/and graphical symbols help children identify better with the teaching activities undertaken in classroom and the curriculum expectations?

An Intelligent Tutor could increase the level of confidence of the child by following Hammond's (2001) scaffolding theory of teaching, increasing the difficulty of lessons to allow children to grasp more complex problems. Could an embodied-pedagogical agent displaying strong positive emotions when the child works on software lead him/her to gain some confidence or help capture its attention on a particular concept it did not understand by expressing emotions such as surprise or dissatisfaction?

1.3 EMOTION, CHILDREN, AND LEARNING

People from all ages, cultures, and backgrounds, express emotions and elicit emotional responses from others in everyday life situations, from greeting someone in the street, to throwing away an object in anger. Similarly, the use of any product can elicit strong emotional responses, which influence the way a person appropriates the product, and its use in the long-term (Holbrook, 1985). The goal of this PhD's work as regards to emotions and affective computing is not to produce a new theory on the impact of emotions for children's learning processes. However, it aims at investigating how the current theories can be applied to OLM applications in order to facilitate learning and increase children's learning, and their motivation to learn.

In a first section, the different theories on children's emotional development and ability to recognize and express emotions will be discussed. To follow, section 1.3.2 will outline the current theories on the interplay between emotions and learning, i.e.: what emotions to consider in educational settings and what are their impacts on learning. Next, a review of the different computational models of emotion will be presented in a third section: How are emotions usually represented and computerized? How are affective components considered? What systems do exist and how could they be adopted to this PhD's work? Section 1.3.4 then will conclude on a description of current research in one branch of the field of affective computing, concerned with capturing the user's feelings and emotional states during software interaction. The final section will conclude on the emotional aspects of learning for child users with a discussion on its applications for intelligent tutoring systems.

1.3.1 CHILD EMOTIONAL DEVELOPMENT THEORIES

Humans feel, perceive, and communicate emotions when confronted with everyday life situations. Such emotions impact a person's evaluation of, and response to a context or a person's state/action, in terms of physical or biological signals sent to the brain, or ignite cognitive processes of emotional responses (Damasio, 1995). Charles Darwin (1998) considered emotions to be a vital element of humans' evolution process. It constitutes a useful resource centre in order for humans to make informed decisions while including the emotional factors. For example, a baby suddenly crying will evoke an emotional response from his/her mother: she will recognize the situation as the child being hunger, hurt, or in need of attention from previous experience, and will act accordingly. Emotions and emotional responses can help gain more awareness of a dangerous situation and decide upon an action to make by comparing previous experiences. It is also a means of communicating with others more efficiently, with a repertoire of body language learned through age and culture (Izard et al,

1993). The posture, facial expressions, and physiological cues of an interlocutor can change his/her discourse, and define a specific emotional state, such as intimidating, comforting, or indifferent.

Charles Darwin is the founder of the first theories on human emotions. From his hypothesis (Darwin 1872) on the innateness and universality of certain facial expressions of emotion, two schools of thought were built:

- Discrete emotion approach, pursued by psychologists such as Ekman, Friesen, Elisworth, or Izard (Ekman & Friesen, 1978; Ekman et al, 1972; Izard 1972; Izard et al, 1993), considering some emotions to be universal and “basic”, and thus pre-determined and hard-wired (Ekman, 1972), while others are considered as “higher cognitive emotions” with a manifestation linked to human’s culture and social codes (LeDoux, 1996).
- Ethnological approach, followed by biologists such as Eibl-Eibesfeldt or Hass (Eibl-Eibesfeldt, 1971, 1973; Haas, 1970; Haas & Harrison, 1977). They consider that to experience a particular emotional state, a person has to learn it from another person.

Salovey et al (Salovey et al, 2000) define the concept of *emotional intelligence* as the capacity to express, understand, and manage your own emotions, as well as perceiving and interpreting those of others. Recognizing and understanding one’s own emotional state and the state of others is important during the learning process. The development of emotions is also affected by cultural variations (Tronick & Morelli, 1991). A connection has been drawn between how people understand emotions, and the development of language and literacy skills (Pons et al, 2003). Children’s ability to learn and speak the ‘language of emotions’ develops through maturation and through social contact with others. could a learning framework including Bruce’s social-learning theory therefore help children understand emotions better and make use of them within an learning context?

“The child’s growth will not be fully understood until we gain greater insights into those aspects of development we call emotional” p.39 Kagan 1978

Izard et al (1993) state that a child’s ability to identify, distinguish and label emotions depends on the particular emotions considered, and this ability improves with age. Their research showed that “basic” emotions such as ‘joy’ were recognized at an earlier stage than other emotions such as ‘anger’ and ‘sadness’. Pons decomposes the development of emotional competence into three age-related stages (Pons et al., 2003). First, four to five-year-old children understand basic emotions, with an average of three (‘happy’, ‘sad’, ‘fear’) out of nine emotional components correctly recognized. An interim group of children (six to nine yrs) were found to understand basic emotions, and also were becoming self-aware and beginning to grasp more complex emotions such as ‘pride’, ‘guilt’ or

‘shame’. The oldest age group, children of ten to eleven years, understood on average eight emotions out of the total nine considered.

Seven principal skills underpin emotional competence and these include emotional awareness; ability to discern and understand other’s emotions; use of a vocabulary of emotion and expression; capacity for empathy and sympathy; differentiating internal emotional and external emotional expression; emotion regulation and coping; and finally, awareness of emotion communication in relationships (Stein et al, 2000). Some studies showed that young children spontaneously talk about their own affective states or other’s emotions with some knowledge of how they felt versus how they *will* feel (Bloom, 1998; Harris, 1997). Whilst young children tend to focus on positive emotions, an early study (Saarni, 1988) found that older pupils reveal their negative emotions, even when aware of the negative consequences that could occur. Young children grasp of people’s emotional reactions differs depending on the beliefs and desires that they have about a situation. Even when two people are confronted by the same objective situation, they may appraise it differently depending on what they want to believe. As a result, a situation that delights someone may upset somebody else, and children can understand that (Harris, 1989).

Pupils first conceive of emotional states and experiences as belonging to one unique emotion. Stein et al (Stein, et al., 2000) examined the development of the awareness of multiple emotions, and determined that children from six years were able to identify feeling several emotions when focusing on one situation. However, emotional valence, an essential notion for representing and understanding both positive and negative emotions, appears to be acquired on average around ten years of age (Harter & Whitesell, 1989). Valence, a bipolar scale ranging from pleasant to unpleasant, is a dimension frequently used in theories of emotion and in the design of affective instruments.

One significant issue with interpreting the foregoing study results relates to whether children are aware of experiencing several emotions simultaneously, but due to their developmental stage are only able to cognitively verbalize it as one (Harter & Buddin, 1987). The ability of children to express mixed feelings – emotional input involving basic emotions of different valence (Izard, 1994) – was studied systematically, by asking children to describe situations that would provoke each of several emotions, two emotions in succession, and two emotions at the same time (Harter, 1983; Harter and Buddin, 1987).

From these experiments, Harter postulates that children develop through several stages:

- From three to six years, children are not able to conceive two emotions begin provoked either successively or simultaneously. They can readily describe situations that will elicit one of the “basic emotions”, but deny that it is possible to provoke two.
- Around six to eight years of age, children begin to describe situations tat will elicit two emotions, but only successively (e.g. “I’d feel excited about getting on the roller coaster, but scare once it got going”). They still doubt that it is possible to experience two feelings at the same time.
- Only when reaching seven to eight years of age do children begin to describe situations that would be likely to cause simultaneously two emotions. They describe situations with two feelings of the same valence, either as a single incident, or two distinct incidents that might occur at the same time.
- The integration of two opposite feelings in a “mixed feelings” situation appears only in children aged ten and more. They first describe situations that involve two separate concurrent situations, before describing a single episode with two intertwined aspects.

This postulate was tested in an experiment (Harris, 1989; experiment 2): two groups of children were asked to describe how a character would feel when experiencing situations of the same valence, and then conflicting situations. Both six and ten year olds considered the character to feel happy in response to positive events, and sad, angry, or scared in response to negative events. When the event included both positive and negative aspects, however, the two age groups offered different replies. The six years olds almost always focused on only one of the two events, while the older children were much more likely to say that the story character would feel a mixture of positive and negative emotions.

Gordon (1986) and Harris (1989) claim that children are aware for their own mental states, and can project them on other people using a mechanism that depends crucially on the imagination. In Harris (1989), an experiment showed that preschool children (four to six year olds) are able to take into account a person’s desires in post-hoc explanations of their judgment about emotions, and also take them into account in making judgment in the first place. Children’s imagination, and more specifically their capacity for make-believe, allows them to entertain possible realities. It therefore leads them to perceive other people’s thoughts and feelings according to their plans, hopes and fears. When capturing feelings or emotions from others, they can then adapt their behaviour or own emotions to the context, using emotional contagion or congruence (Harris, 1989). Could the use of an EPA displaying emotions influence children’s expectations and motivation to learn by capturing feelings of positive reinforcements, or critical views on their success?

Children can imagine the beliefs and desires that other people may have, even if they themselves do not share them. They can use such make-believe premises to reach new conclusions (Dias and Harris, 1988). In (Harris, 1989), an argument was made that children understand other people's mental states by relying on a distinctive type of imaginary understanding. Children are aware of their own mental states, and can project them on to other people using a mechanism that depends crucially on the imagination, and includes four concepts:

- Self-awareness: children know when they want something, or expect something, when they have made a mistake, or feel sad (Johnson, 1988).
- The capacity for pretence: young children have a powerful imagination. From about the age of 18 months, children start to endow physical objects with pretend properties, and thereby create a make-believe situation.
- Distinguishing reality from pretence: although children occasionally mix the world of reality and the world of pretence, they do not show any systematic confusion between the two worlds. They can construct make-believe worlds, yet it is not confused with the real world. The pretend desires, beliefs and emotions with which the child endows physical objects with pretend properties and imaginary characters are not confused with real desires, belief and emotions.
- Desires, beliefs and emotions: the ability to pretend allows children to engage in an imaginative understanding of other people's mental states. Given their capacity for pretend play, children can imagine wanting something they do not actually want.

1.3.2 INTERPLAY BETWEEN EMOTIONS AND LEARNING

The significance of emotions experienced in educational settings has been recognized by researchers in different fields, including personality research that has analyzed students' test anxiety for decades (Zeidner, 1998), research on achievement motivation (Heckhausen, 1991), and more recent educational studies focusing on a variety of emotions in education. The three most commonly used frameworks on the link between learning and emotion will be introduced in this section: Pekrun et al's *control-value theory of achievement emotion* (2002a, 2002b); Linnenbrink & Pintrich's *achievement goal theory* (2002a, 2002b, 2003, 2004); and the *appraisal theories*: the *OCC Model* (Ortony, Clore & Collins, 1988), and Kort et al's *four quadrant model* (2001).

CONTROL-VALUE THEORY OF ACHIEVEMENT EMOTION

The control-value theory of achievement emotions (Pekrun, 2000) is based on the premise that current approaches to “achievement emotions” share a number of common basic assumptions, and can be regarded as being complementary rather than mutually exclusive.

Achievement emotions can be grouped according to their valence (positive vs. negative; or pleasant vs. unpleasant), and to the degree of activation implied (activating vs. deactivating) (Linnenbrink, 2007), and object focus (activity vs. prospective outcome vs. retrospective outcome). A three-dimensional taxonomy was defined in (Pekrun et al, 2002a) using these dimensions.

Table I.3: Control-Value Theory of achievement emotions

Object Focus	Appraisals Value	Control	Achievement Emotion
Outcome/Prospective	Positive (Success)	High Medium Low	Anticipatory joy Hope Hopelessness
	Negative (Failure)	High Medium Low	Anticipatory relief Anxiety Hopelessness
Outcome/Retrospective	Positive (Success)	Irrelevant Self Other	Joy Pride Gratitude
	Negative (Failure)	Irrelevant Self Other	Sadness Shame Anger
Activity	Positive	High	Enjoyment
	Negative	High	Anger
	Positive/Negative	Low	Frustration
	None	High/Low	Boredom

In the theory, achievement emotions are defined as emotions tied directly to achievement activities or achievement outcomes. Achievement can be defined simply as the quality of activities or their outcomes as evaluated by some standard of excellence (Heckhausen, 1991). By implication, most emotions pertaining to student’s academic learning and achievement are seen as achievement emotions, since they relate to behaviours and outcomes that are typically judged according to standards of quality – by students themselves and by others. The second series of emotions experienced in education are social emotions, for example a student’s caring for a friend in the classroom. Achievement and social emotions can overlap, as in emotions directed towards the achievement of others (contempt, envy, empathy, admiration instigated by the success or failure of others, see Weiner, 2007).

The theory also addresses the effects of achievement emotions on student’s academic engagement and performance. Specifically, it is posited that emotions influence cognitive resources, motivation, use of strategies, and self-regulation vs. external regulation of learning. In line with research from self-efficacy theory, Pekrun reported high correlations between self-efficacy in

mathematics certain types of emotions: joy (+.62), anger(-.53), anxiety(-.63), hopelessness (-.67), and boredom (-.42). These emotions were also associated with values of achievement in maths but the associations were much weaker than those with self-efficacy.

The results of the implementation of Pekrun et al's (2007) theory in experimental design show that:

- Emotions help focus attention on the object of the emotion. Positive emotions relating to the activity were assumed to focus attention on the activity, thus benefitting performance. On the contrary, the use of emotional display when facing failure can help distract users from the overall activity and therefore increase the cognitive resources available for concentrating on the task itself.
- Positive activating emotions such as enjoyment of learning are assumed to increase interest and strengthen motivation. Negative deactivating emotions such as hopelessness or boredom, are held to be detrimental for motivation.

PEDAGOGICAL GOAL AND EMOTION RECOGNITION

Boekaerts (2007) assumed that students have two priorities in the classroom, namely:

1. Increasing their assets by improving their competence in areas they value
2. Keeping their well being within reasonable bounds.

Students who appraise a learning activity favourably – meaning that positive cognitions and emotions are dominant – start activity in the mastery or growth pathway. The assumption is that the former students perceive the task as congruent with their personal goals values and needs, and therefore the learning activity is energized from the top down. By contrast, students who perceive environmental or internal cues during the learning process that signal a mismatch between the learning activity and their personal goals, needs and interest, experience negative cognitions and emotions, which prompt them to switch to the well-being route.

Clearly structured, cognitive activating material and challenging task demands that match student's abilities likely benefit student's competences and interest, thus positively affecting their appraisals and emotions. When the level of achievement or understanding of children is set too high, another emotion can be aroused: *boredom*. This is in line with Piaget's disequilibrium theory, where emotions are integral to belief change, in that some emotions tend to emerge during mismatches between one's goals and/or beliefs and teacher's perceptions of what is occurring (1974).

Graham, Hudley, and Williams (1992) pointed out that some students have the inner strength to tolerate negative emotions and hide them while they are trying to make a mental representation of an ambiguous situation and design a plan of action. Other students, particularly those who want to attain a performance goal (e.g. I want to be better than my peers on this task), may not be prepared to

tolerate negative affect associated with the difficulty of the task. In addition, they may not be prepared to pay the mental (e.g. effort) and social costs of admitting difficulty (embarrassment), (see Butler, 1992).

Students who pursue recognition and reward goals might give up more easily than students who pursue mastery goals when negative emotions are elicited during the learning process. As Pintrich (2000) explained, a performance goal creates a mindset that prompts students to interpret negative emotions as signals that the task is too complex to look smart in front of the class. By contrast, students who want to master a difficult task because they value its content have created a mindset that helps them to interpret negative emotions as signals that more effort needs to be invested to achieve a learning goal or that social support needs to be solicited.

ACHIEVEMENT GOAL THEORY

In the control-value theory, only “emotions” are considered as part of the emotional spectrum. Barrett & Russell’s affective circumplex model (1998) considers mood states as well as emotions, and categorizes affect based on appraisal processes, which highlight the dimensions of object focus and valence, as illustrated in Figure I.2.

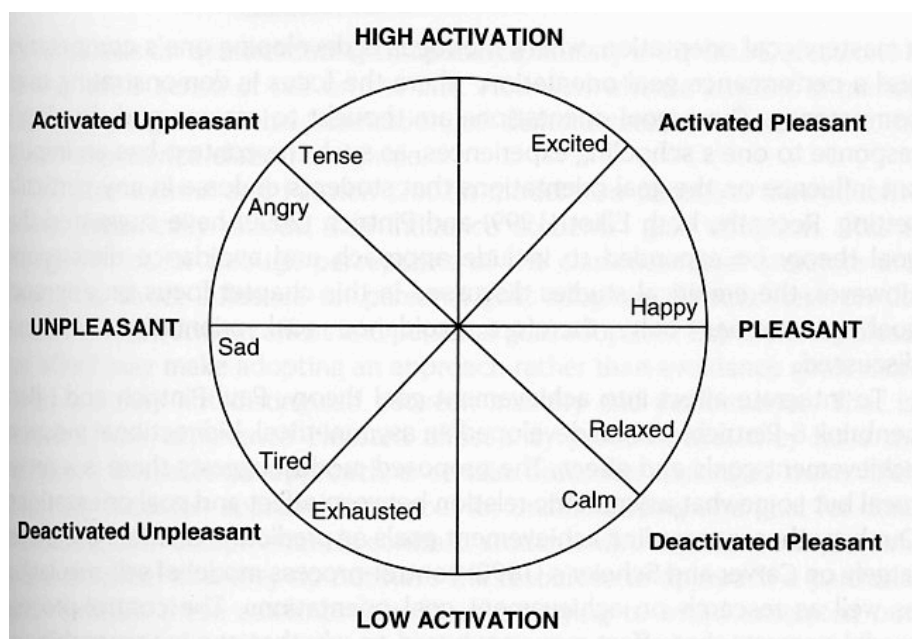


Figure I.2: Affect circumplex (Barrett & Russell, 1998)

Russell’s model emphasizes the bipolarity of affect, which is consistent with historical measurements of mood (Lorr et al, 1982; Meddis, 1972). In this circumplex model, pleasant (e.g. happy, satisfied) and unpleasant (e.g. sad, gloomy) states occupy opposite ends of the same dimension, thus creating a bipolar continuum of positive versus negative feelings.

From this circumplex model of affect, Linnenbrink & Pintrich (2002a, 2002b, 2003, 2004) have developed a general model for educational settings integrating research on achievement goal theory, affective states, and school engagement. According to achievement goal theory, goal orientations provide a framework for interpreting and reacting to events (Dweck & Loggett, 1988), with two primary goal orientations: A *mastery goal orientation*, where students' focus is to develop his/her competences and knowledge acquisition; and a *performance goal orientation*, where students' focus is to show/demonstrate his/her skills.

The model supposes a link between affect and goal orientations, based on Carver & Scheier's (1990) control-process model of self-regulation: Affect is supposedly based on how close a student is to reach a particular goal, and his/her progress rate towards that goal. Results from studies performed at primary, secondary school, and university level found mastery-approach goal orientations to be positively related to pleasant affect and negatively related to unpleasant affect (Linnenbrink, 2007).

Attribution theory proposes that student's motivation to learn is directly rooted in their beliefs about why they succeed or fail at tasks (Weiner 1972). If students can be taught to alter these beliefs, for instance to understand that failure is the result of a lack of effort instead of a lack of ability, then their motivation to learn and learning outcomes can be significantly improved (Robertson, 2000). For example, (Woolf et al, 2010):

- *General attribution* messages encourage students to reflect about math learning in general
- *Effort attribution* messages reinforce the idea that effort is a necessary by-product of learning, and are specially tailored to situations where students are investing effort but are struggling;
- *No-effort attribution* messages are more empathic than effort attributes and are designed to help students realize that effort is necessary to learn, and generated when students are not investing effort;
- *Incorrect attribution* messages are generated to motivate students after they provide an incorrect response, by re-formulating how they perceived errors.

APPRAISAL THEORIES OF EMOTION

Appraisal theories of emotion represent both one of the most ancient, as well as one of the most recent theoretical traditions in affect: Ancient, as leading philosophers such as Aristotle (384-322 BC.), Descartes (1596-1650), or Spinoza (1632-1677) assumed that the major emotions are differentiated by the type of evaluation or judgment a person makes of the elicited event. Most recent, as active theoretical development was achieved only in the early 1980s.

As defined by Izard,

“Emotions tend to occur in patterns or sequences, sometimes with rapid shifts in emotion signals and emotion-related behaviours.” p9, (Izard, 1982)

The theory (Weiner, 1972; Scherer & Peper, 2001) assumes an emotion architecture that is based on people’s subjective evaluation of appraisal of the significance of events for their well-being and goal achievement (Lazarus, 1991). It highlights the dynamic character of the emotion process (an episode rather than a state), as well as the organized nature of the emotion components.

The OCC Model

Most research in design, educational technology, and HCI adopts the cognitive or functional model of emotions developed by Ortony, Clore and Collins (1988). The model focuses on the cognitive or situational structure of emotions as implied by emotion concept, and follows the appraisal views of emotion. They postulate that all aspects of specific emotions, including their characteristic response patterning, is dictated by the psychologically significant situation the recipient of the emotion is involved in.

This model states that emotions serve an adaptive purpose; they are the mechanisms that signal when events are favourable or harmful (Frijda, 1986). The model focuses on three aspects of the world: events, agents and objects. People focus on events for their consequences, on agents for their actions, and on objects because of our interest in their properties. The current focus of research has been on positive emotions which has some impact on children’s motivation to learn. However, negative emotions are also part of learning and can be helpful in ‘learning how to learn’ (Kort et al, 2001): induced by failure, they can focus on the child’s attention on the issues encountered and avoid the repetition of errors.

The OCC model addresses the problem of representing emotions not by using sets of basic emotions, or by using an explicitly dimensioned space, but by grouping emotions according to cognitive eliciting conditions. In particular, it assumes that emotions arise from valence (positive or negative) reactions to situations consisting of events, agents, and objects. With this structure, Ortony, Clore and Collins outlined specifications for 22 emotion types, as given in the boxes in appendix A.

Additionally, they included a rule-based system for the generation of these emotion types. The underlying structures for different emotions are seen as an analogy to grammar in language, a rule system that describes speech and emotion, but that is not consciously evoked when speaking or emotion.

Four Quadrant Model

The four quadrant computational model brings together the constructivism and appraisal theories: its structure follows Russell & Barrett's circumplex of affect (1998), but emotions users feel during learning are viewed as an evolutionary process, interactive, and elicited through an appraisal process.

Kort et al's (2001) Four Quadrant Model of emotions for learning is designed as a circumplex model of affect, according to two axes: the *Emotion axis*, valence of the emotions in terms of pleasantness, and the *Learning axis*, representing the construction of knowledge upward ("learning") and the discarding of misconceptions downwards ("un-learning"). It is divided in four quadrants according to the degree of pleasantness and learning acquired, as illustrated in Figure I.3:

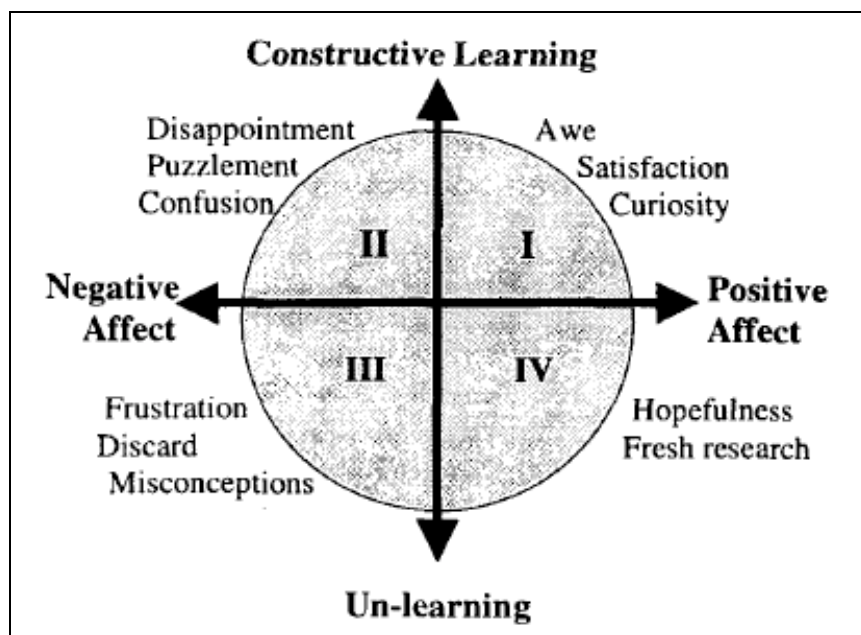


Figure I.3: Four Quadrants Model of Emotion

The model aims at being used with a third dimension, time, where the evolution of the user's emotional state can be followed in the progression from quadrant to quadrant.

The following intervention strategies were defined from video-taped observations of six to eleven years old users working on two educational applications: *Incredible Machine*, and *Gizmos and Gadgets* (Kort et al, 2001):

- When the learner is in quadrant I, little to no intervention is needed as s/he is happily engaged in exploratory and/or discovery learning.
- When the learner is in quadrant II and has recognized/acknowledge his/her work is erroneous or incomplete, the intervention focuses on providing the emotional support required to survive and emerge from the negative emotions that may arise during the retreat and recovery phase of the learning cycle.
- Intervention in quadrant III is the most challenging and uncertain, as it aims to successfully “grieve the loss and get on with life” (Kort et al, 2001).
- In quadrant IV, the learner has gone back to a good path of learning construction, and there has been an improvement of the understanding of the subject at hand. The intervention required should include some scaffolding, and several ways of supporting the student can be applied.
- Finally, the user returns to quadrant I with a fresh insight and new ideas, the intervention should be celebrating and praising, reinforcing the feeling of pleasure and delight that accompany successful learning, for the student to want to engage in another loop.

The main particularity of this model, when compared to other circumplex, is that emotional states during learning are viewed as a dynamic and evolutionary process: children’s emotions are categorized into quadrants, and their emotions go from one quadrant to another through time. This is in line with the notion of emotional scaffolding, induced by children’s reactions to the learning material and the role of teacher-student interaction in classroom teaching. Teachers can take the students through an evolution of their emotional states, in order to maximize their motivation, attention on the task at hand, and learning.

EMOTIONAL SCAFFOLDING

Expert teachers are very adept at recognizing and addressing the emotional state of learners and, based upon their observation they take some action that positively impacts learning (Kort et al, 2001). Expanding on traditional views of scaffolding, Meyer and Turner (2007) define emotional scaffolding as temporary but reliable teacher-initiated interactions. They support student's positive emotional experiences to achieve a variety of classroom goals. For example, emotional scaffolding may sustain and enhance student's understanding, motivation, collaboration, participation, and emotional well-being. It differs from other forms of positive teacher-student interactions as the support aims to increase student achievement and autonomy in a particular development competency, and often in several areas simultaneously.

The extent to which emotional upsets can interfere with mental life is no news to teachers. Students who are anxious, angry, or depressed don't learn; people who are caught in these states do not take in information efficiently or deal with it well.

- Daniel Goleman, Emotional Intelligence (1995, p2)

Within instructional interactions, scaffolding has been identified as temporary teacher support to achieve two interrelated goals: 1) provide support only as necessary, and 2) move from a position of shared responsibility to one in which the student takes ownership (Wood et al, 1976). It is this balance between teacher support and student autonomy that has linked the scaffolding metaphor to Vygotsky's (1987) theoretical construct of "zone of proximal development" defined as an interpersonal space within which a teacher provides support as needed while negotiating a gradual transfer of responsibility to students.

An important feature of classroom that has been associated with positive affect is an emphasis on learning goals in contrast to performance goals. Learning goals promote effort and developing competence in comparison to others (Ames, 1992). Teachers initiate and sustain learning goals by scaffolding 'positive affective classroom climates' through instructional interactions.

According to the theory of emotional scaffolding, children go through an evolution of their emotional state while learning. When picked up by teachers, the learning goals and tasks are chosen taking in this factor, in order to maximise emotional scaffolding. Should learning take place using technological devices, it is therefore interesting to investigate ways to capture children's emotional state, and to follow its evolution.

1.3.3 CAPTURING YOUNG USERS' EMOTIONAL STATES.

Research into product design has clearly demonstrated that products can elicit strong emotional responses. Positive emotional responses to user interface (UI) designs promote further use of the product whilst negative emotions often indicate the presence of usability problems. Web-based products and services and their subsequent usage levels provide a prime illustration of the need for good product design. Past evaluation of UI designs centred on performance indicators. However, the current trend human computer interaction is on developing methods and tools that can measure the quality of the user experience.

Clearly, emotional responses influence the way a person appropriates a product, and its use in the long-term (Holbrook, 1985). Accordingly, when users are confronted with new educational software, they react emotionally to application features and content. The emotions elicited (e.g. pleasure, satisfaction or dissatisfaction) can vary for users on their first use related to the attractiveness of the software, or after long-term use where users focus more on feature and content usage, than interface aesthetics.

This evolution of emotions can also be seen to affect the way in which children interact with educational software products, with positive emotions increasing motivation to both interact and learn. The use of any product results in users feeling and expressing emotions in a depth and complexity that depends on their involvement in an activity, the type and success of the activity performed, and also on the users' ability to capture, understand, interpret and express emotions. In the field of educational technology and games, there is an interest in non-intrusive methods that help capture the quality of the user's experience with the application. Non-intrusive methods need to be carefully designed so as not to adversely affect any positive emotional responses the software is evoking.

There is a tendency, perhaps because these examples are easy to think of, to imagine the worst forms of emotions in computers (Picard, 1997). For example, consider if you repeatedly typed the wrong input to your machine, and the machine finally threw insults at you, or refused to let you do anything anymore in a state of frustration or desire for revenge from keyboard abuse. Clearly, there must be more to benefit from affective computing than this.

In this section, a review of the different methods to report children's emotional experience while using technology is reported. Affective user feedback can be divided into two categories: explicit and implicit. Implicit methods sense the behaviour of the user and deduce an emotion. Explicit methods ask the use to input affective feedback directly. Eliciting emotions using 'intrusive' methods (by adding artefacts or physiological sensors) has been undertaken (Alsmeyer et al, 2007; Heraz & Frasson, 2008; Isbister, 2006; Isbister et al , 2006). However, such intrusive technologies are difficult

for educators to use in class, disturb the learning process and teaching strategies (Van Dam et al, 2005), distract children, or must be used in experimental settings. The aim of this PhD's work is to enable investigations as to children's use of OLM applications within classroom settings, using readily available material, in order to enable the deployment of software for long-term use. This therefore excludes any type of intrusive sensing method, such as readings of heart rate, brain waves, or skin conductance currently used in affective computing research (Woolf et al, 2010b).

EKMAN'S FACIAL RECOGNITION SYSTEM

The first method presented in this section is part of the implicit user-feedback category. Ekman et al (1976) developed a measurement system based on components of emotion expressions. The Facial Affect Scoring Technique (FAST) highlights appearance changes in these areas of the face for six emotions: joy, surprise, sadness, disgust, anger, and fear. Coders code each of three areas of the face separately, with the rest of the face blocked from view. Formulas are then applied to the data to identify emotion expressions. Ekman and Friesen (1978) later replaced FAST with the Facial Action Figure Coding System (FACS). FACS is an anatomically based, comprehensive system that codes all observable facial actions. FACS contains twenty-four discrete action units and twenty miscellaneous facial actions, as illustrated in Figure I.4. It identifies the same six emotions that FAST does. The illustrative material for FACS consists of (voluntary) muscle contractions modelled by the authors and other adults.

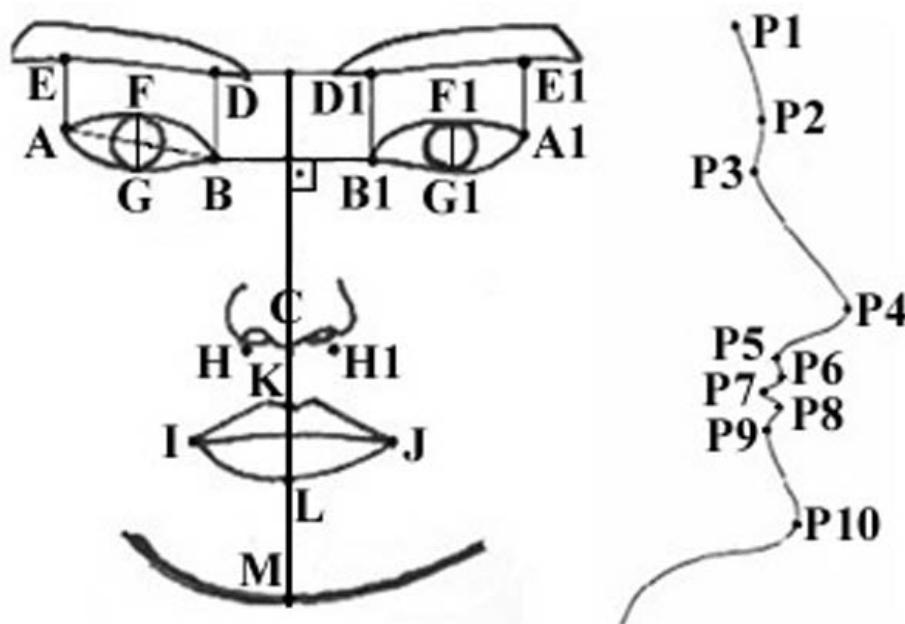


Figure I.4: FACS facial reading points

The FACS system distinguishes between 44 so-called Action Units (AUs). Take Action Unit 12, for instance, which codifies a smile. This is defined as a rising of the corners of the mouth, accompanied by a horizontal stretching of the mouth. The main issue with this system is its complexity, which leads to a very time-consuming tool. It takes up to a hundred hours to learn all the rules, and after that it takes an hour to interpret twenty-four images, which is equivalent to just a single second of motion pictures.

This system has been extended to automatic visual recognition of affect in faces, using figure detectors and a high-end artificial intelligence rule-based programming core, ISFER, in the form of a neural network of previously encountered pictures (Pantic & Rothkrantz, 1999). These detectors perform automatic spatial sampling of the features in digitised facial images. One problem is that the computer is unable to recognise the facial muscles and action units directly from a digital image without previously having analysed a view of the same face that shows a neutral facial expression (expressionless face). Currently, the automatic system manages to correctly assess 89 out of 100 images.

Automatic affect recognition is known to work adequately in laboratory environments, but currently fails when used as a generic mechanism to capture affect due to the large number of confounding factors in this particular context (Zeng et al, 2009). Furthermore, this method only takes into account the emotions children wish to portray to others, but not necessarily what they think of their emotional state. Harris' work (1989), described in details in section 1.3.2, showed that children, by the age of six, are aware of their own emotions, and the social impact of portraying such emotions to others. They will be able to hide, exaggerate, or refrain from experiencing an emotion according to the context, and their understanding of the consequences of such emotional display.

When capturing children's emotional state, the interest for learning may lie in the ability and willingness of children to give some personal feedback on their experience. In order to capture such affective intention, an explicit method of affective user feedback therefore seems more appropriate. Furthermore, as the age of the children concerned with this PhD's work ranges between seven to eleven years, some of the youngest children might not be able to read well. Therefore, the review of the methods will concentrate on the graphical self-reporting theories and associated tools.

SELF-ASSESSMENT MANIKIN

The *Self-Assessment Manikin*, also called *SAM* (Lang, 1985; Bradley et al, 1994), is a graphical instrument that measures emotions according to three dimensions (PAD): Pleasure (i.e.: valence), Arousal, and Dominance (i.e.: control). An emotion is represented with a combination of one element taken from each dimension, without requiring language, as illustrated in Figure I.5.

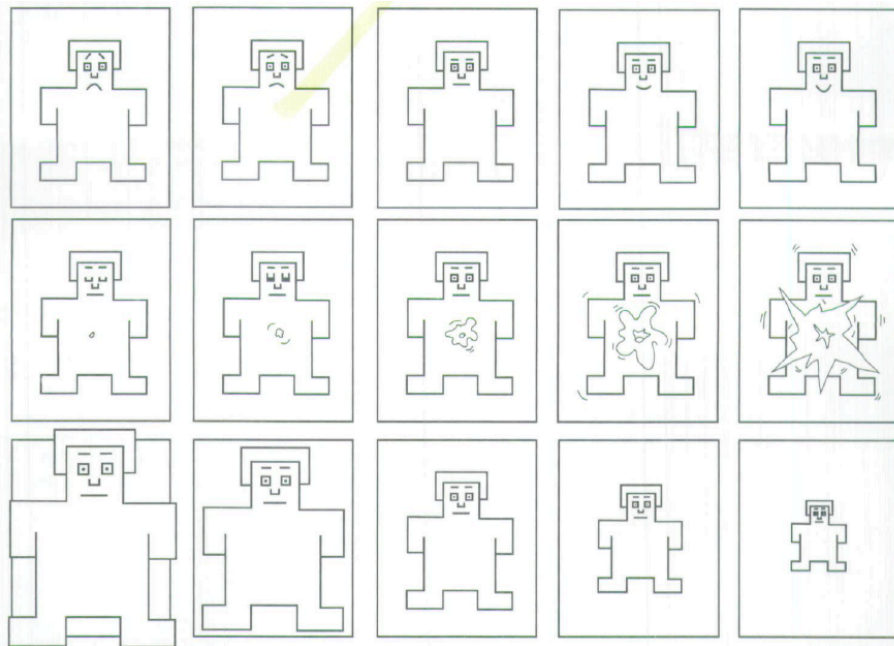


Figure I.5: Self-Assessment Manikin: the three PAD dimensions

SAM ranges from a smiling, happy figure to a frowning, unhappy figure when representing the pleasure dimension. For arousal, *SAM* ranges from an excited, wide-eyed figure to a relaxed, sleepy one. For dominance, *SAM* ranges from a large figure (in control) to a small figure (completely controlled).

Mehrabian and Russell (1977) demonstrated that all emotions could be accurately described in terms of three independent and bipolar dimensions: pleasure-displeasure, degree of arousal, and dominance-submissiveness. An emotional state is therefore described as “a state that can be described as a region within a three-dimensional space” (Mehrabian and Russell, 1977). The *SAM* method has been shown to require low cognitive involvement and be more attractive than verbal self-report methods for adolescents and adult users (Lang, 1985). This instrument is a culture-free, language-free measurement, suitable for use in different countries and cultures (Bradley et al, 1994). However, it requires the users to understand the three dimensions of the PAD system in order to rate their emotions with the method.

AFFECTBUTTON

The *AffectButton* (Broekens & Brinkman, 2009) is a measurement device to input mixed and graded intensity affective feedback. Traditional self-report methods to capture emotions ask a user to input discrete emotions, such as Ekman's six universal emotions (Ekman et al, 1976), with or without intensity. The particularity of the *AffectButton* lies in the use of continuous feedback on core affective dimensions, the PAD dimensions developed in the previous section. It involves a low cognitive effort from the users due to the dimensions chosen, and does not need the users to understand the three dimensions of PAD, unlike the *SAM* method.

The user can select a large range of affective values from the whole PAD spectrum by selecting one point of the space on the button. The affective values are represented by the rendered facial expressions, interpolated between nine prototypical expressions, as illustrated in Figure I.6.

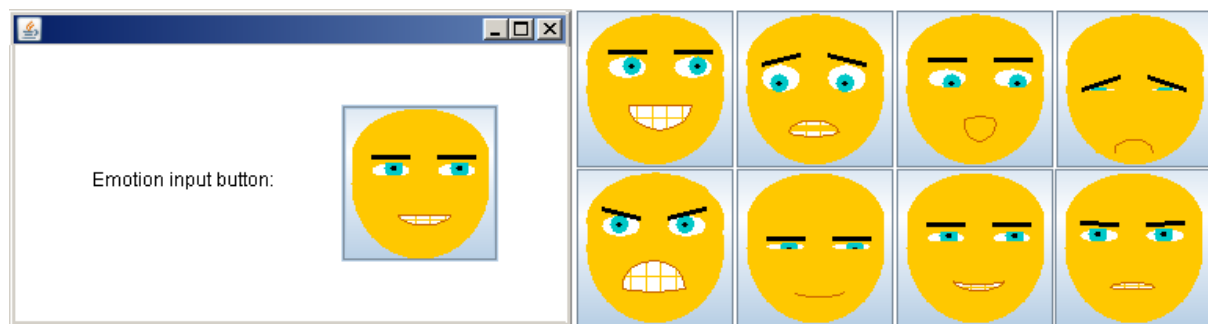


Figure I.6: AffectButton and the nine prototypical expressions (elated, afraid, surprised, sad, angry, relaxed, content, frustrated, and neutral).

According to Harris (1989), children initially learn about the situation that provoke basic emotions like happiness or anger, and then begin to understand the situations that provoke more complex emotions like pride and guilt. In doing so, a simplifying assumption has been made: that any given situation elicits a single emotion. Although this may often be the case, it is not always true. When we are about to start some new venture, we often feel a mixture of apprehension and excitement. Conversely, when an opportunity slips through our fingers, we might feel disappointment tinged with relief. These cases of mixed emotions are intriguing because they involve ambivalence: a combination of positive and negative emotion. This is the main drawback of the *AffectButton* and *SAM* methods: the tools do not allow the expression of 'mixed-feelings', which hold value in affective research on learning. For this reason, a new method of self-report, following the theories hereby presented, is needed.

1.3.5 EMOTIONS FOR LEARNING: DISCUSSION

In this section, evidence from research studies show that emotions have a role to play in the learning process. It also highlights that children's understanding of, and experience with emotions differ as they age. The issues raised from research in children's emotional development, emotions for learning, and how to capture a child's emotional state are now discussed, in terms of the implications for the design of OLM tutors.

The most commonly used method to include affect into ITS in research is through embedded pedagogical agents (EPA): "visually represented, computer-generated characters in pedagogical roles, such as visual instructors, mentors, or learning companions" (Haake & Gulz, 2009).

One issue with affective EPAs is the definition of the agent's behaviour towards user-software interaction, in order to maximize students' motivation and learning (Woolf et al, 2010b). Several theories of emotions have been integrated into ITS, with results on children's motivation and learning process:

- Providing empathy or support strongly correlates with learning (Graham and Weiner 1996; Zimmerman 2000) and the presence of someone who cares, or at least appears to care, can be motivating. Researchers have developed systems that are able to control and influence the learner's emotions, using empathy. For instance, the "*affective companion*" adapts to the learner's emotions, by adjusting the difficulty of an exercise (Isen 2000); whereas the "*affective tutor*", on the other hand, is itself affectively adaptive to user's emotions (Estrada et al, 1994). McQuiggan developed and tested the advantage of affect-based empathic responses on hints and suggestions (McQuiggan and Lester 2006). In the *Wayang Tutor* (Woolf et al, 2010; Arroyo et al, 2010), the learning companions are empathic in that they visually reflect the last emotion reported by the student, as long as this emotion is not negative, e.g. companions do not mirror frustration or boredom, and give feedback corresponding to the Attribution theory described in section 1.3.3.
- Bringing together the emotional self-regulation, and Kort's theory of emotions in learning, Burleson and Picard (2007) implement an approach that uses affective agents in the role of peer learning companions to help learners develop meta-cognitive skills such as affective self-awareness for dealing with failure and frustration.

Some issues are therefore raised concerning the choice of emotions to display, their integration into pedagogical activities, and their impact on learning:

Children's understanding of emotions grows as they age. What emotions would therefore be suitable for children aged seven to eleven? Should there be a unique model of emotions, understood by all? Or rather a more extensive sample of emotions, where the young children have an adaptive access to the model, as their understanding grows?

While children aged seven to eleven have some understanding of the valence of an emotion, they will represent an emotional state either by a unique emotion, a succession of unique emotions, or a combination of basic emotions using "mixed-feelings". What system of representation of emotional states could be used in educational environment for children to express their feelings? Should affective feedback be represented by a single emotion, or more complex emotional states?

The majority of research studies favour emotions of positive valence as software feedback when considering learning. However, some theories such as Mandler's cognitive disequilibrium theory (1984) or Kort et al's four quadrants theory (2001) show the importance of considering negative in the process. Are positive emotions the only ones children can benefit from in a learning environment, or could the display of *frustration*, *boredom*, or even *disappointment* also be useful for learning?

Emotional scaffolding is used by teachers in classrooms, and has the potential to sustain and enhance children's learning, motivation, and emotional well being. EPAs have the potential to digitally act as a teacher, or a coach, within ITS. Could the use of an EPA create an emotional attachment from the children, and help them view the activities more like a 'game', increasing their motivation to play with software? Could such agents help children's emotional scaffolding process by expressing strategy of emotional display similar to classroom teachers' emotions?

A second set of issues concerns the digital representation of emotions, for children to understand and recognize emotional display. The "*Oz project*" (Bates, 1994) has been interested in making agents believable, by giving them the illusion of life. It contains a variety of synthetic characters that may not look like any real creatures, but that are designed to be able to powerfully influence their audience as if they were real. In the magical *Disney Animation* (Thomas and Johnson, 1981), the Disney masters emphasize the importance of each character having a clear emotional state at all times. They describe numerous techniques for accomplishing this, arguing about the portrayal of emotions is what gives the Disney characters the "illusion of life".

How can emotions be digitally represented in such a way that children aged seven to eleven recognize them?

What are the long-term effects of using emotional pedagogical agents? Are the short-term learning effects persistent for exposure over an extended period of time? How does the user's perception of the agent change over time and what are its effects on the child's motivation to learn?

From such observations, one can wonder whether using affective embodied pedagogical agents in OLM tutors might help children gain more understanding of the learning environment.

Could affective feedback embedded in digital characters, at the centre of the child's make believe and pretend world when playing on software, help them to engage more in the software activities and be more motivated to learn?

Could the desires, beliefs, and emotions portrayed by the characters lead children to understand better the pedagogical goals given, and help them in structuring their knowledge acquisition via the use of the LM information?

1.4 OPEN-LEARNER MODELING ENVIRONMENTS

Intelligent tutoring systems (ITS) are computer-based learning environments structured around four main components: the domain knowledge (also called expert system), the task environment, the student model (or learner model), and the pedagogical module. Open-Learner Modelling (OLM) environments are ITS applications where the learner model is partly accessible for inspection/editing by different actors within the learning process (learner, teachers, peers, parents, etc...). These systems have been found beneficial for learning and motivation to learn, as well as for children's reflection on their own knowledge (Zapata-Riviera, 2003). However, issues remain, in particular with user's interaction with, and understanding of the learner model component.

In section 1.3.1, the basics of ITS will be presented, along with the different technological revolutions that changed certain aspects of ITS research. To follow, the OLMs will be introduced, with a presentation of their advantages for learning, as well as the issues still remaining to be studied. Finally, ways to represent and interact with the learner model will be investigated, in particular through the use of pedagogical agents to act as the user-software interaction medium.

1.4.1 USING TECHNOLOGY TO TEACH MATHEMATICS IN CLASSROOMS

Walker et al (1994) stated that the use of computers can provide more active learning, less mental drudgery and are better tailored to individual learning rates. In the last decade, new technologies have been increasingly used in order to support people in their everyday lives. This applies to the domains of education and pedagogy, and led to significant changes in teaching and learning practices (Van Dam et al, 2005).

"IT's role is to augment (not to replace) the teacher, to provide human-centred tools that encourage and support adaptability and flexibility, and to enable appropriate modes of learning", (Van Dam et al, 2005), p30.

Computer-based learning environments seem to be of great value in their capacity of producing personalized learning environments in addition to increasing children's motivation in learning. However, a significant number of such existent systems do not suit the teachers' educational needs or technological abilities (Van Dam et al, 2005).

Straker et al (1989) define three ways the computer can be used within the classroom: as a teacher aid, a learning resource, or a tool for the child to use in doing pedagogical tasks. The key to developing mathematical knowledge is to form a base of understanding and to provide the means for applying it. Hopkins et al (Hopkins et al., 1996) define three key skills that are needed for children to

effectively tackle unfamiliar processes in mathematics: mathematical decision making, communication and reasoning mathematically.

One type of instructional strategy implemented in the form of an Intelligent Tutoring Systems is *drill-and-practice* applications. It promotes the acquisition of knowledge or skill through repetitive practice, and therefore help learners master materials at their own pace (Chassé & Lefebvre, 2006). This type of applications is usually used as a reinforcement tool: the learners have already some knowledge of the specific concepts/skills to acquire, and practice such skills at their own pace on software through the repetitive drills. By practicing on a specific list of concepts, considered as essential for future learning, the learner can build skills that become the building blocks for more meaningful learning (Chassé & Lefebvre, 2006).

One of the advantages of using a drill-and-practice application rather than teachers providing exercises in classrooms is the diversity and possible personalisation of the drills. On the one hand, the use of randomization in the building of the drills enables the creation of drills that are similar in principle and structure, but not in content, much more easily than a human can do by hand (Chassé & Lefebvre, 2006). On the other hand, by having some knowledge of what the child is able to do, and has already done on software, through the learner model component, drill-and-practice applications could produce drills that focus on the learner's weaknesses and current learning. The instantaneous feedback given to the learners after performing on software could also help them focus on how well they are doing through praising, increasing their motivation to learn, and the particular misconceptions or lack of knowledge they might have, leading them to revise in a more effective manner.

The use of scenarios of application of knowledge could also help them to develop decision-making skills and thus the ability to interlink mathematical concepts and deepen their understanding. The child could acquire communication skills by interacting with the computer-based tutor. In accordance with Bruner's findings (1966) the practice on a huge number of relevant examples and scenarios to apply mathematical knowledge could accelerate the child's ability for mathematical reasoning, or at least build the basis of reasoning skills acquisition. Finally, the use of animations, emotive characters, sound and colours could also help convey ideas and concepts to children, helping them focus their thinking and practice using mathematical language appropriately (Shneiderman, 2004).

With interactive resources feeding children's natural curiosity and enthusiasm for interaction, the introduction of computers in the teaching techniques is a valuable addition and an effective way for the teacher to capture the attention of the child. It should also provide students with questions to demonstrate the use of concepts in different contexts. The tutor finally must be a coach, allowing the child to ask questions, summarizing the material that has been learnt and giving the child feedback on its progress. In order for those tools to be appropriately used in classroom, to what extent can open

learner-modelling techniques be used in the design of the tutors? Could the insertion of affect in parts of the tools design help teachers and children access/interact with the information they are provided with?

1.4.2 INTELLIGENT TUTORING SYSTEMS

Providing a personal training assistant for each learner is beyond the budget for most organisations, including schools participating in this PhD's investigations. However, a 'virtual' training assistant that captures the subject matter and teaching expertise of experienced trainers provides a new option. The goal of Intelligent Tutoring Systems (ITS) is to provide the benefits of one-to-one instruction automatically and cost-effectively. Research in the field of computer-assisted instruction (CAI) has been described under numerous names but all refer to computer software that aims at instructing. ITS and adaptive hypermedia systems add the "intelligence" to it.

Polson & Richardson (1988) defined the aim of an ITS as to "develop a simulation of human problem solving in a domain in which the knowledge is decomposed into meaningful, humanlike components and deployed in a human manner".

As detailed in (Corbett et al, 1990) and discussed by Polson and Richardson (1988), any intelligent tutoring system is composed of four different components, illustrated in figure I.6:

- Task environment: defines the learning activities in which the students can engage;
- Domain knowledge (or *expert system*): heart of an ITS, it provides the bases for interpreting student actions by representing the content knowledge that the student is acquiring;
- Student model (or *learner model*, *LM*): represents a record of the student's knowledge state composed of two distinct parts: an overlay of the expert system (copy in which each knowledge unit is associated to an estimate of how well the student has learned it), and a bug catalogue (set of misconceptions or incorrect values each representing whether the learner has acquired this misconception in one of the sessions);
- Pedagogical module: responsible for structuring the instructional information by sequencing topics to ensure an appropriate structure of pre-requisites (Capell & Donnenberg, 1993) and advise students on the different learning activities to ensure they understand the material (Corbett & Anderson, 1995).

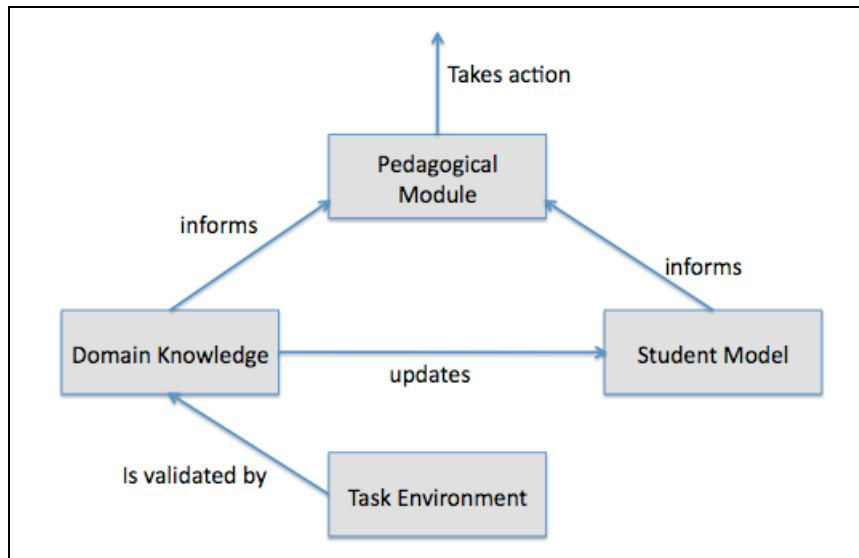


Figure I.7: Traditional ITS architecture

The student-tutor interaction proceeds as follows: Students engage in one of the various learning activities present in the task environment. Students' actions are then evaluated with respects to the domain knowledge component, and a decision is reached as to the value of the interaction in terms of learning. From this evaluation, an update of the student model is performed to incorporate the last pedagogical actions and achievements. Finally, the pedagogical module delivers instructional actions based on the information provided by the domain knowledge, and the current student model. A more detailed structure of the tutor instruction process can be found in Appendix A, extracted from (Shute & Psotka, 1996).

A student model is the key to provide an intelligent interaction with the student. Understanding the student's behaviour, attitudes, and state of knowledge, making use of different diagnosis techniques and sometimes with the collaboration of students and/or teachers, is a key factor to be considered in the design and implementation of any system that cares about the student. Student models can be divided into three major categories: bandwidth, target knowledge type, and differences between the expert's and the student's knowledge. According to their bandwidth (i.e. approximate mental states, intermediate states, and final states) it is possible to determine which algorithm can be used for diagnosis within the model. The second category, target knowledge type, includes procedural knowledge, declarative knowledge, and qualitative mental models of complex systems. The final category, differences between the expert's and the student's knowledge, covers overlay student models, bug libraries, and bug part libraries.

By studying student modelling research, Kass (1989) was able to identify three major techniques that can be used to relate the student model to the expert model. The first is overlay modelling (Carr & Goldstein 1977). It assumes that the student's knowledge is a subset of the expert model. The overlay modelling explains all differences between the student's behaviour and the

expert's behaviour as lack of skills by the student. Second, differential modelling (Brown et al. 1975) considers two kinds of knowledge (namely, what the student should know and what the student could not be expected to know). Finally, perturbation modelling (Brown & Burton 1978, Sleeman & Smith 1981, Johnson & Soloway 1984) represents beliefs beyond the expert model. Bugs, misconceptions and lack of knowledge are considered as perturbations. The perturbation modelling technique assumes that the student and the expert models can be compared.

Self (1999) describes the concept of an interaction model in education as an extended student model that focuses on the interactive process. It addresses aspects, such as: timing of the learner's actions, the contexts in which they occur, and the learner's cognitive state at the time. This interaction model is accompanied by a situation model and an affordance model in a new conceptual architecture of ITSs. Self (1988) identified six major roles for the student model:

- *Corrective*: to help eradicate bugs in the student's knowledge;
- *Elaborative*: to help correct 'incomplete' student knowledge;
- *Strategic*: to help initiate significant changes in the tutorial strategies other than the tactical above decisions;
- *Diagnostic*: to help diagnose bugs in the student's knowledge;
- *Predictive*: to help determine the student's likely response to tutorial actions;
- and *Evaluative*: to help assess the student of the ITS.

McCalla et al. (2000) propose the concept of active learner modelling. In this approach, learner models computed dynamically and 'just in time' support the needs for information about learners in distributed learning environments. "Model" from this perspective is seen as a 'verb', a continuous process of assembling and summarising fragmented learner information from potentially diverse sources. These definitions cover most of the elements that have characterised student-modelling research. Student models have been considered as important component of any ITS.

Brusilovsky (1994) proposes a student model centred approach for intelligent learning environments (ILE) inspired by the use of student models in traditional ITSs and further enhanced in what he calls an advanced, student-model centred, open architecture for ILE. His work has two main goals: first, to build an adaptive ILE where all the components can adapt dynamically to the changing student model, and second, to provide a single representation of the student's knowledge that can be used by all the components of an ILE in order to facilitate continuity. A new student-model centred architecture, based on the use of local views of the model per each module called projections, was proposed. In this architecture students and modules within the ILE could control individual

components separately without interfering with other parts of the system and without losing important information. Barnard and Sandberg describe a general model for learning environments that is broad in scope, has the learner as an important focus, and offers a functional view. These three aspects are essential when representing and comparing a wide range of learning environments.

An approach that has contributed to reduce student-modelling complexity is the use of simulated students or learning companions. In this approach, the student is represented by an agent companion or simulated student, which is created to simulate the student's level of knowledge of a particular domain. The learning companion approximates a typical student and in that sense represents a form of active student model (Chan & Baskin 1990; Chan 1996). Once the agent companion is running, it can be used to act as a peer, virtual tutor or tutee. Could the insertion of a pedagogical agent in traditional OLM architecture help users in the understanding and acquisition of the information provided by the model?

1.4.2 OPENING THE MODEL TO STUDENTS: ADVANTAGES AND ISSUES

Part of the learner modelling research community studies user models considering the user as a learner and using the models in the design of educational software. The learner model is the foundation of system with the potential to treat learners as individuals. Research in user-modelling for computer-assisted instruction systems gives an increasing interest in the Open Learner Models (OLM), models of the user that are available for viewing by the learner and sometimes others. These systems have been found to be of help in improving the learner's performance and their understanding of the learner model component (Baker et al, 2004; Luckin & Hammerton, 2002) as well as promoting reflection (Bull & Pain, 1995; Zapata-Riviera, 2003).

OPENING THE LM TO WHOM?

The inspection of, and the interaction with, the user model have been realised with a variety of methods extending from allowing a student to view his/her progress using a skill meter (Corbett & Bhatnagar, 1997) to a process of negotiation between the learner and the system to define the content and representation of the model (Bull & Pain, 1995; Dimitrova, 2003). Learner models have also been externalised not only to learners themselves but also to peers, instructors/teachers, and even parents according to the purposes of openness (Bull & Kay, 2005). Open learner models for children need to be simple in format in order that the children can understand them. Examples include knowledge level represented as coloured magic wands for 7-8 yrs olds (Bull et al, 2005) or smiling faces to represent the different levels of knowledge for 8-9 yrs olds (Bull et McKay, 2004).

Learner models opened to instructors have been used to allow the instructor to follow the evolution of a student's knowledge (Vincent, et al., 2005); to help instructors adapt their teaching to

the individual or the group (Grigoriadou et al 2003; Yacef, 2005); or to help instructors organise learning groups (Müehenbrock et al, 1998). The use of stereotypes in *PepiStereo* (Vincent, et al., 2005) or meta-cognitive tools in *Exploring the Nardoo* (Hedberg et al, 1994) offer teachers the opportunity to individualise instruction, by having access to information of great pedagogical value concerning their students. However, the lack of teacher's motivation in using ITS has proven to be correlated to their lack of experience in using new technologies, as well as the gap between the pedagogical content of software and the teacher's expectations.

HOW TO DESIGN, COMPARE, AND ANALYSE OLM APPLICATIONS?

Definition of the purpose of the modelling

An ITS can be classified according to the purpose of the modelling. Kok (1991) proposes a classification under eight dimensions, with four headings: “why”, “who”, “what”, and “how”.

- *Why are users modelled?* What is the overall aim of modelling users? Which parts of the system need user information?
- *Who is modelled?* What is the role of the user being modelled in relation to the system? How individual are the user models, which and how many users are represented by a model?
- *What is modelled:* What aspects of the user are meant to be represented by the user model contents? What is contained in the user models, and what are the interpretation methods of the contents?
- *How are users modelled:* What are the methodologies of modelling and what are its sources? Which user modelling techniques are used?

This classification helps inform the design by asking the questions necessary for future OLM construction, and proposes a method to follow to research information in the literature as to each part of the application construction. However, the research questions investigated in the OLM community cannot be restricted only to why, who, what, and how. In particular, a deeper understanding of the *why* is needed to enable a proper understanding of the application purpose, and how the learner model can be designed, accessed, and edited.

Definition of the dimensions and content of the LM:

Bull and Kay (Bull & Kay, 2007; Bull & Pain, 1995; Bull et al, 2005) then proposed a method to describe and analyse OLM: The SMIL[©] Open Learner Modelling Framework. This framework can be used to compare different OLM systems emphasizing on the differences in access, content, and edition of the learner model. Similarly, it can also be used in design to select the purposes and specifications of the application to design, and to check the reality of the LM built against the initial purposes of the design.

In the SMIL[©] framework, seven purposes of user-model openness have been defined: accuracy, reflection, plan/monitor, collaboration/competition, navigation, right of access/control of/trust, and assessment.

They were associated to research aims as follow in (Bull & Kay, 2010):

- *accuracy, reflection, plan/monitor*: Promoting meta-cognitive activities such as reflection, planning and self-monitoring;
- *reflection, plan/monitor*: Allowing the learner to take greater control and responsibility over their learning; encouraging learner independence;
- *collaboration/competition*: Prompting or supporting collaborative and/or competitive interactions amongst groups of students;
- *collaboration/competition*: Facilitating interaction between learners and peers, teachers, and parents;
- *navigation*: Facilitating navigation to materials, exercises, problems or tasks, etc... where links are available from the learner model;
- *accuracy, right of access/control of /trust*: Increasing the accuracy of the learner model data if the user is allowed to contribute additional or corrective information, to enable a more precise adaptive interaction to follow;
- *right of access/control of /trust*: Increasing learner trust in an adaptive educational environment by showing the system's inferences about their knowledge;
- *right of access/control of /trust*: The (non-educational) issue of people having the right to access electronic data about themselves.

- *assessment*: Supporting assessment – in particular providing formative assessment opportunities for students, but also enabling the learner model to be used as a summative assessment;

Some of these research aims and purposes have been investigated in other types of applications, and are not specific to OLM software. For example, assessment, collaboration and competition, and navigation are considered in all ITS systems. However some the purposes are specific to the nature of an OLM such as accuracy, right of access to the model, control of the model content, and trust in the learner model, and lead to the specific research questions and issues:

Self (1999) argues that the learner model provides an opportunity for learners to take some responsibility for its content, thus improving its accuracy. As the learning process is an evolving process (cf. section 1.2), it is difficult for OLM applications to accurately represent the learner's current knowledge. Indeed, a learner may have forgotten skills since the last time s/he used software, performed more study outside of the system therefore increasing substantially his/her knowledge base from one software use to another, or might receive help during the task which might over-evaluate the system's belief of his/her capabilities. Therefore, some systems, with edition or negotiation rights, enable the learners to express their own belief over their current knowledge, as well as the system's. The measure of accuracy, later investigated in (Kerly & Bull, 2006), represents how much the system's and the child's beliefs about their knowledge become identical, after using the application: it quantifies the number of topics that were found in disagreement in the pre-test, and then were agreed on in the post-test. The higher this number reveals to be, the better the model is deemed "accurate".

In OLMs, risks could result from learner control over their model. For example, the learner may not give an accurate self-assessment when editing model content, and so provide incorrect information to their LM (Kay, 2001). Indeed, learners do not always view themselves as bearer of the actual skills they master: some can over-estimate their capabilities, while others may under-estimate it. Therefore should the user be authorized to edit his learner model directly, by supplementing the system's beliefs about the content, or should other means of interaction be provided to compare and contrast the child's and the system's beliefs?

Defining the type of interaction with the LM content

In order to define the type of access, control, and edition of the information of the LM, there is a need to classify the different ways in which learners can access the LM content, or representations of its underlying content. Different types of OLM applications have been defined and researched, differing in the level of access and control over the learner model (Bull & Kay, 2010):

- *Inspectable learner models*: the model is available for user viewing, but with no additional interaction possible. In such applications, the learner model is completely under control of the system. Most of the OLM applications are from this category, and address the requirements for people to be allowed access to electronic data about themselves, aim at raising learner awareness of their knowledge, and prompt reflection, planning and formative assessment.
- *Editable learner models*: available for user viewing, where the learner is able to change, or edit its contents. In Flexi-OLM (Mabbott and Bull, 2006), the system may offer evidence or information to demonstrate where it disagrees with the learner's viewpoint. The learner can override this if they wish, and so ensure that their desired changes to the model are effected.
- *Persuasive or Negotiated learner models*: allow an intermediate level of learner control, where the user and system are both giving information for the learner model. The information given by each actor of the learning process can be different, and/or complementary. The communication between the actors does not necessarily need to end up as one single assessment of a specific skill, but can represent different opinions in the form of views (later referred to as system's view and child's view of the learner model). Some of the communication is within a form of *persuasion*, where the learner aims to change the model data by showing to the system that their own assessment of their skills is accurate (they are usually asked to perform some kind of test to affirm their knowledge); to a truly balanced distribution of the control of the learner model contents between the user and the system by a *negotiation process* aiming at forming an 'agreeing' learner model. Should the system and the user not arrive at an agreement, the conflicts of representation are kept in the model. Such systems vary from a mixed control over the learner model (Dimitrova 2003) to full user control (Ahmad and Bull, 2008).

In Mitrovic and Martin (2007), university level students performing on an OLM application, the SQL-tutor, were found to perform significantly better than their peers, working on an application without access to their model information (named Closed-Learner Model application).

It seems that some learners may trust a system to infer their knowledge more than they trust themselves to assess it or change it. Trust in the accuracy of the learner model may even be more important if learners have no control over its content (Ahmad and Bull, 2008). For this reason, could a system offering children the possibility to express their view of the learning achievement, and/or negotiate the system's assessment of such knowledge, help children make a better use of LM information for learning?

NEGOTIATION WITHIN OLM SYSTEMS

It has been suggested that students may be less comfortable with editing the model; they may prefer an OLM that offers less direct control, such as a negotiation process (Mabbott et Bull, 2006). However the experiment was performed on only 8 university students, and for a very limited amount of time. While the student's preferences were investigated, the potential for better learning of the negotiation process still remains to be clearly defined.

Negotiation describes any communication process between individuals that is intended to reach a compromise or agreement to the satisfaction of both parties (Nierenberg, 1977). Negotiation involves examining the facts of a situation, exposing the both the common and opposing interests of the parties involved, and bargaining to resolve as many issues as possible.

The dominant framework for understanding fundamental negotiation strategy is the dual concerns model (Rubin, Pruitt, and Kim 1994). The model postulates that individuals in negotiation have two somewhat independent concerns: a concern for realizing one's own substantive outcomes in the negotiation, and a concern for helping the other party achieve their outcomes, usually in order to strengthen a positive working relationship with the other party.

The strength of one's concerns on each of these two dimensions dictates one of five major strategies, as described in (Lewicki et al. 2003) :

- *Contending* (i.e. competing or dominating) is the strategy to consider when one has a strong concern for one's own outcome and has little concern about the other's outcomes. Negotiators employing a contending strategy try to obtain the best outcome possible only for themselves.
- *Yielding* (i.e. accommodating or obliging) strategy consists in negotiators showing little interest in attaining their own outcomes but strongly care that the other party achieves their goals, possibly in order to build a stronger future relationship with the other party.

- *Inaction* (i.e. avoiding) occurs when negotiators have little interest in achieving either their own outcomes or the other's outcomes and is equal to retreating or withdrawing from the negotiation.
- *Collaborating* (i.e. problem solving): When negotiators show high concern for attaining their own outcomes and a high concern for whether the other party attains his or her outcomes, they pursue a collaborative or problem-solving strategy, in order to maximize their joint outcome and to reach a 'win-win' situation.
- *Compromising*: negotiators pursue a compromising strategy when they display a moderate effort to pursue their own outcomes and a similarly moderate amount of effort to help the other party achieve his or her outcomes and strengthen the relationship.

Since parties negotiate largely to enhance their own outcomes, the two most common strategic approaches are competing (contending) and collaborating (problem solving). Thus, the dominant choice confronting the negotiator is whether the relationship with the other party is important or not.

Negotiation within OLM applications can be undertaken in a variety of ways, for example: by menu selection (Bull and Pain, 1995), dialogue games (Dimitrova 2003), and chatbots (Kerly and Bull, 2008). Laboratory trials (Bull and Pain, 1995; Dimitrova, 2003) have shown the potential of menu selection and dialogue games for engaging learner reflection and enhancing the accuracy of the learner model. However, the negotiation methods were found restrictive and somewhat unnatural. Further research in OLM and negotiation introduced artificial agents as an interaction means with the learner model and the user, facilitating the negotiation process by a natural interaction with the agent (Kerly and Bull, 2008). It was found a more flexible way to express the learner's view in a naturalistic and intuitive way. On the other hand, this method requires a high level of reading and writing skills which is not necessarily reached at the developmental stage of our intended learner's age group (who can be as young as seven years old). Therefore, research into ways to facilitate the interaction between the learners and the model content within the negotiation process are still open to investigations.

In order to support the negotiation process functionality, the learner model must store distinct records of the learner's and the system's beliefs about the learner's knowledge (Kerly and Bull, 2008). Two separate measures are usually used, and taken into account in providing adaptive interactions. Studies performed on the Mr. Collins system, in laboratory, (Bull and Pain, 1995), found that students were interested in being able to see the contents of their learner model. According to Bruce's theory of free-flow (2001), the fact to enable children aged seven through eleven to give their opinion on their learning achievements and compare it to the system's beliefs should enable them able to wallow in the learning, therefore increasing motivation and learning gains.

Current systems employ different mechanisms to present and support human interactions with the learner models. But can it be expected from young children to learn and reflect on their knowledge by browsing freely through the model? Should the system detect potential conflicts between the learner's and the system's view of the model and point them out for the students to investigate further? Could a protocol be created for learners to follow through the negotiation process and what would the consequences be on children's learning, motivation, and interaction with LM content?

Whatever technique used for the negotiation process, the key point is that the system and the user have equal rights to initiate and end discussions in negotiated learner models; and have the same negotiation moves available (e.g. offer information, request information, confirm, accept, critique, challenge, refute, justify). Giving the user the opportunity to engage in a negotiation process for the building the LM content has the potential to reinforce the use of such strategies. Could the use of a pedagogical agent to help with the negotiation process influence children's behaviour towards negotiating the model content in terms of strategies employed? What could be the potential benefit for learning of using said strategies in the construction of the LM?

REFLECTION IN OLM SYSTEMS

Negotiation in OLM applications is closely linked to the purpose of reflection: in order to engage in a negotiation process with the system, one need to have reflected on one's own learning achievements, skills, and future learning prospects. In most of the studies performed on OLM, the targeted users are adults, from university students to elders, who can be expected to understand the role of reflection in learning. The educational and developmental benefits of using OLM with child-users to improve reflective processes has yet to be clearly defined, and to date, mixed results have been found as to the willingness of children to use the learner model information, and how they use it. While Zapata-Riviera and Greer (2004) argue that children aged ten to thirteen can perform self-assessment and undertake reflection on their knowledge in association with an OLM, Barnard and Sandberg's study found that secondary school children did not look at their learner model when it was available to them for voluntary use (1996).

Pedagogical approaches developed from Vygostkian theories emphasize the importance of reflection within the learning process. Students are able to assumed increased responsibility for their learning through a heightened awareness of both personal and shared meanings, as they move towards becoming "experts" within the classroom community (Irroyo et al, 2010). Interaction is a stimulus for reflection and plays an essential role in the environment, that involve learners in a discussion about their learning achievements and skill representations (Dimitrova, 2003). When learners are engaged in a discussion/negotiation process about the learner models they are reflecting upon their domain knowledge re-calling and re-considering ideas of which they are aware. This reflective process necessitates users making connections between what they know and the LM representation, and

provides opportunities for the learners to be engaged actively in what they are learning (Bound et al, 1996). By engaging in reflective discussions, students and teachers become aware of the choices and responsibilities within the processes of learning and teaching. Could the use of an OLM system, by mirroring the classroom reflective actions through a process of negotiating the LM content, help children develop their identity as learners and express their views (Cazden, 1988)?

Researchers have explored different strategies to achieve this goal. Strategies used vary in terms of the degree of intrusion and sophistication. For example, Kay (1999) makes available the student model and provides tools to support scrutability, integration of evidence and conflict resolution. In Bull et al. (1999) students are required to provide self-assessment, give and receive feedback from a peer, and explore the system's beliefs. Bull & Pain (1995) explore students negotiating and discussing the student model with the system. In Morales et al. (2001) students use a computer-based simulator and a special graphical interface to interact with a student model composed of a set of rules. The student model is updated based on the student's behaviour while using the simulator. In Dimitrova et al. (2001) students are guided through a constructive dialogue process in which the system can switch between different diagnosis tactics: a discussion between the user and the system takes place with the system providing different tactics to the learner to choose from to explain his learning shortcomings or achievements, and the model is built step by step as the strategies are explored. Although these researchers use a variety of strategies, most of them report that students explored the student model and evidence of student reflection was found (Kerly & Bull, 2006; Zapata-Riviera & Greer, 2004). However, the impact of such reflective process on learning and motivation are not clearly demonstrated, and present such methods as a potential for learning rather than beneficial for learning.

Dillenbourg (1996) states that the mere existence of reflection tools does not imply that students reflect on their learning experiences. Similarly, in promoting self-explanations, Aleven and Koedinger (2000) report that using a simple menu driven graphical interface in which students could explain their answers was not effective. In fact, Aleven and Koedinger state that instructional guidance or some form of tutoring dialogue mechanism should be provided in order to promote self-explanation (2001). Zapata-Riviera's study with child-users (2003) showed that using several guidance mechanisms (i.e. following a protocol, interacting with human peers and artificial guiding agents, interacting with the teacher, and exploring the model as part of a group) to support student interaction with the student model results in students becoming engaged in a reflection process that involves activities such as looking for more information about the domain content using various media, interacting with human classmates or artificial agents in order to learn more about some topic, asking the teacher for an explanation, or defending their position using verbal or written explanations (2002).

To what extent can opening the externalisation of the user-model to child-users create learning environments more appropriate for use by teachers in classroom settings: which representation of user-models are more usable and beneficial? How can they use this information to learn mathematical concepts and increase their level of reflection upon the learning process and their mental representation of learning acquisition? How can OLM systems be incorporated into conventional learning systems? How much control should children have over their learning session and LM information inspection/edition in order for them to engage more and have a better learning experience?

OLM applications seem to be able to encourage user's reflection upon one particular concept to learn. However, one has to wonder how beneficial this reflective process is for children's learning. As discussed earlier in section 1.2.4, reflection can sometimes be harmful for learning. Indeed, a child confronted to a rehearsal of stressful or problematic experience without any resolution offered by software could damage his/her self-esteem and drive away his/her will to learn this particular concept ever again. Could the use of affective feedback help children focus on the positive sides of learning and avoid entering negative loops by increasing their understanding of mistakes made and giving them resolution opportunities? Could the use of an affective pedagogical agent to show how concepts are related and what the child knows, and therefore also facilitate resolutions?

1.4.3 REPRESENTATION OF AND INTERACTING WITH THE MODEL: USING AFFECTIVE EPAS AS AN INTERACTION MEDIUM.

One area of OLM research is concerned with how to represent parts of the learner model for the users to access/inspect/manipulate. The views of such OLM can vary in terms of content: overview of a knowledge level, organization in structural shapes between concepts, illustration of misconceptions linked to the concepts they are dealing with, ... The second aspect dealing with LM representation is the tools used in order to visualize learner model information, for example pedagogical targets (Brusilovsky and Sosnovsky, 2005); liquid level (Papaniklaou et al, 2003), smiley faces (Bull & McKay, 2004); or tree growth (Lee & Bull, 2008).

The tools range from basic Likert scales to Bayesian representations in 3D graphs, as illustrated in Figure I.8:

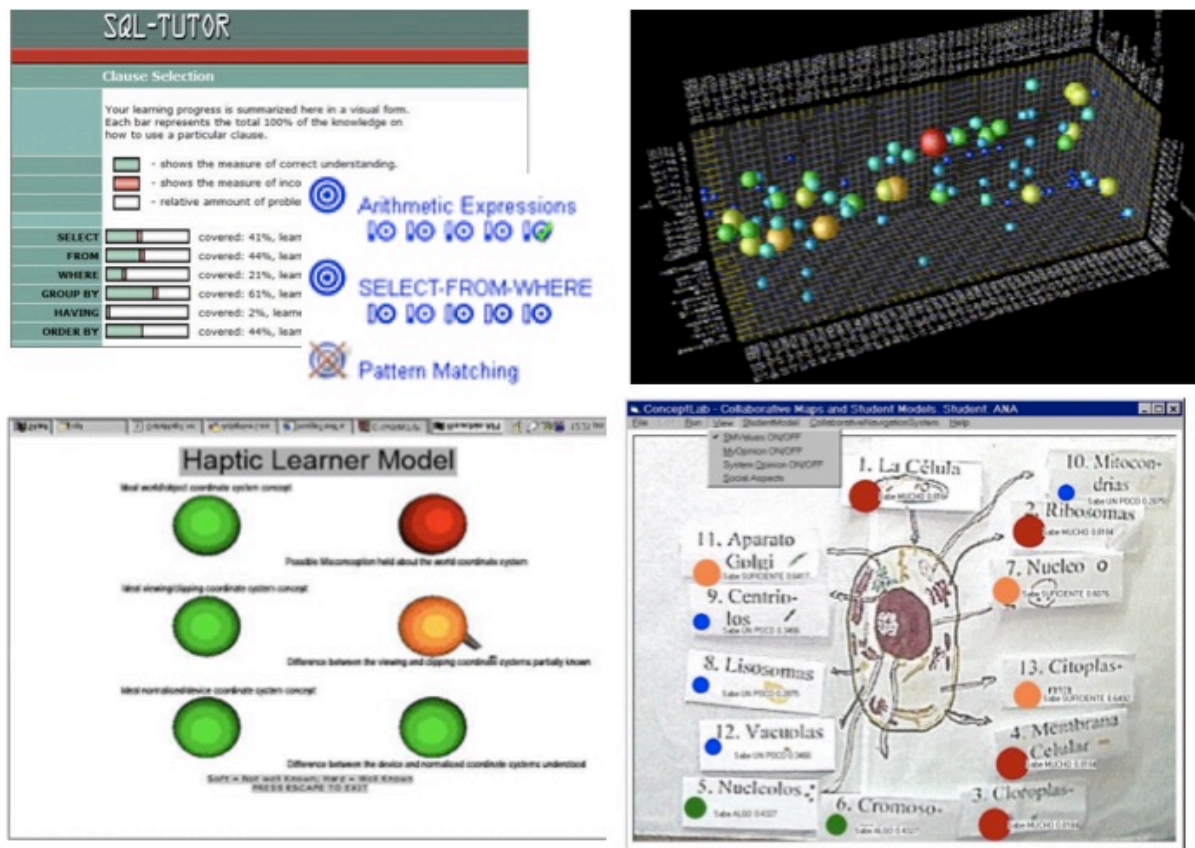


Figure I.8: Different graphical representations of LM (Mitrovic & Martin, 2002), (Brusilovsky & Sosnovsky 2005), (Mazza & Dimitrova, 2004), (Lloyd & Bull, 2006), (Zapata, 2003)

Four types of LM views are most commonly used in OLM systems, as described in (Bull & Kay, 2010):

- Lecture view (lists topics according to the order they are presented in the activities of the software),
- Related concepts view (logical, hierarchical structured grouping of subject matter),
- Concept map view (represent the conceptual relationships between concepts),
- and Pre-requisites view (shows a suggested order for studying topics).

In Flexi-OLM (Mabbott & Bull, 2004), those views are available for the learners to choose from, and visualize at will.

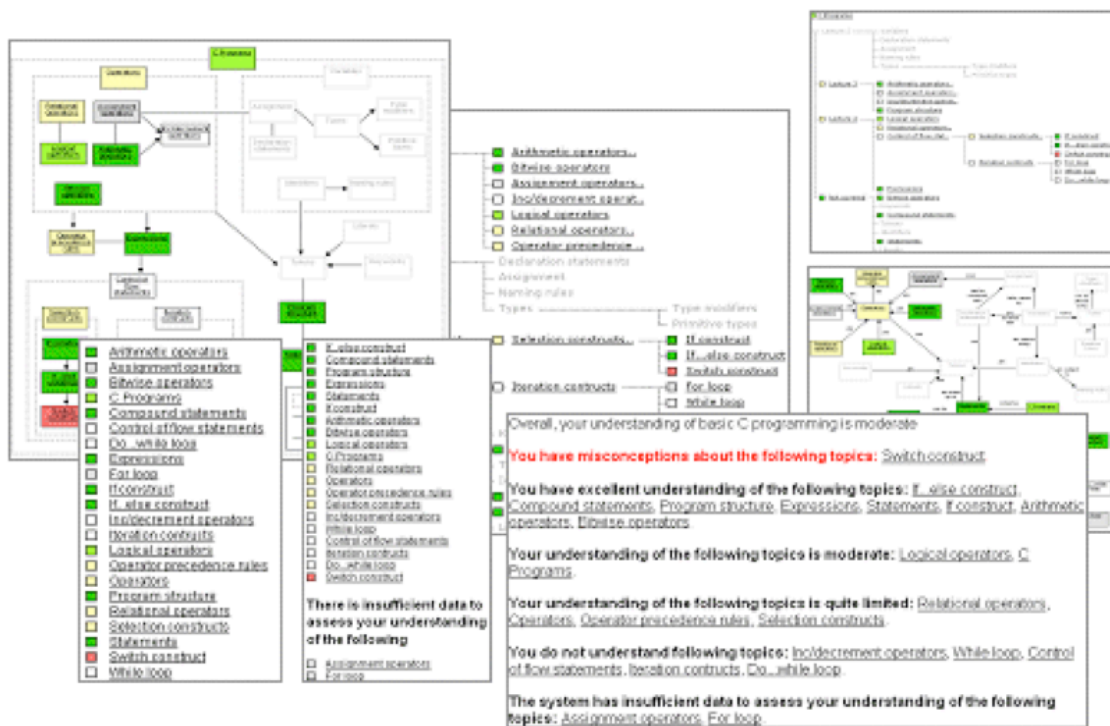


Figure I.9: Views of the LM, in (Mabbott & Bull, 2004)

Each learner does not integrate and visualize knowledge in the same way, which brings forward the potential of using different visualization and modelling techniques in order to suit every learner's needs: if too many techniques are used, system maintenance and update becomes very difficult to manage for designers, but also for teachers (when producing new material, one has to think of its adaptability for each visualization proposed).

Bull & Kay (2010) define one of the main research issues in OLM systems as the learner's ability to understand the information of the LM, what to make of such information, and the possible outcomes in terms of learning of LM manipulation. This is particularly true for child users as they are still evolving in their mental representation of concepts, ability to relate concepts to one another, and reflect upon their own learning.

In Mabbott & Bull's study (2004), university students have been found to possess clear individual preferences when multiple detailed views of the model were available. However, the study failed to show any of the views as more useful than others. One can wonder whether the same tendencies apply for child users, and which type of learner model representation might be better understood by this specific type of learners. The issues of lack of use, or non-effective use of the learner model representations provided, were reported in Kay (1995) and Barnard & Sandberg (1996) in OLM systems for adult users. Providing a representation that children find more useful and easier to understand could help counter such problems. Bull & Mabbott's study (2006) on students' preferences in learner model representation tools highlighted skill-meters as the most commonly used tool in OLM systems. It also insisted on the usefulness of providing users with additional visualization tools, such as graphs (second most preferred in their study).

Colour is a common way of indicating knowledge level in both simple and more complex open learner models. Adaptive link annotations as found in many adaptive educational hypermedia environments are a simple form of open learner model using, for example, coloured bullet points in a contents list to indicate a learner's readiness to attempt a topic (Weber & Brusilovsky, 2001; Weber et al., 2001). Colour has also been used in concept maps and hierarchical knowledge structures to indicate knowledge level (Mabbott & Bull, 2006); and text labels in combination with the size of the text (Uther & Kay, 2003; Kay & Lum, 2003; Kay & Li, 2006). Systems may provide explanations of how the learner model information was used in adaptation (Czarkowski et al., 2005), or where the evidence in the learner model came from (Kay, 1997; Kay & Lum, 2005). Some systems also allow the learner to look back at previous knowledge states or anticipated future states (Bull et al., 1995).

Which types of visualisation/representation of this user-model are beneficial for children's understanding and how does it affect their learning accomplishments and motivation? Are traditional visual techniques such as Skill-Meters or Graphs enough to incite children to use the information, or could pedagogical agents help gain more insight from children on the information presented in the views by serving as interaction medium between the model and the child-users?

Embodied Pedagogical Agents (EPAs) can be found in educational software in increasing numbers, and under different visual representations, different communication styles, and adopting one or more different pedagogical roles, within the same application. EPAs are usually embedded within the software to aid social and communicative features, (Baylor, 2005; McQuiggan et al, 2008). They

are used in computer-assisted learning applications for users ranging from children to elderly people, to help them in software navigation, or in learning content or development of meta-cognitive skills (Yee & Bailenson, 2007; Baylor, 2005).

Recent research tends to confirm the “Persona Effect”, which states that learning is facilitated by a life-like persona that expresses affect (Wood & Wood, 1999). Pedagogical agents are considered as “social interfaces” (Lester et al, 2000) that can act as social actors within educational software,. They are believed to convey information or feedback to learners in a more intuitive and engaging way in order to positively affect learner attitudes, motivational state and even learning gains (Craig et al., 2002). Shneiderman (1998) argues that interface personas can make users feel inferior if their mode of instruction is too patronising. Thus, the pedagogical agents should not be intrusive to the functionality of the system, and its affect on the usability of the system as a whole must be carefully planned and considered.

However, studies showed that visual appearance, communicative style, and pedagogical roles of agents impact their acceptance and trust, and change the way people interact with them (Nowak & Rauh, 2008; Yee & Bailenson, 2007; Haake & Gulz, 2009; Baylor & Kim, 2005). For example, the use of vocal feedback in addition to textual one in the EPAs has been found to help university level students to learn better than when confronted with an EPA producing textual feedback only (Craig et al, 2002; Atkinson, 2002).

Pedagogical agents implementing various forms of adaptive instruction use their own view of the student model to keep track of student progress. Some of these pedagogical agents can implement some form of collaboration or negotiated assessment using a view of the student model to support formative dialogue between students and teachers (Kerly & Bull, 2006; Zapata-Riviera, 2003). Zapata & Greer’s (2004) work on inspecting Bayesian learner models through the use of guiding artificial agents showed the potential for agents to facilitate student interactions with the model: As illustrated in Figure I.10, they were found to help users navigate through the model and find conflicting nodes, and increase the levels of student reflection.

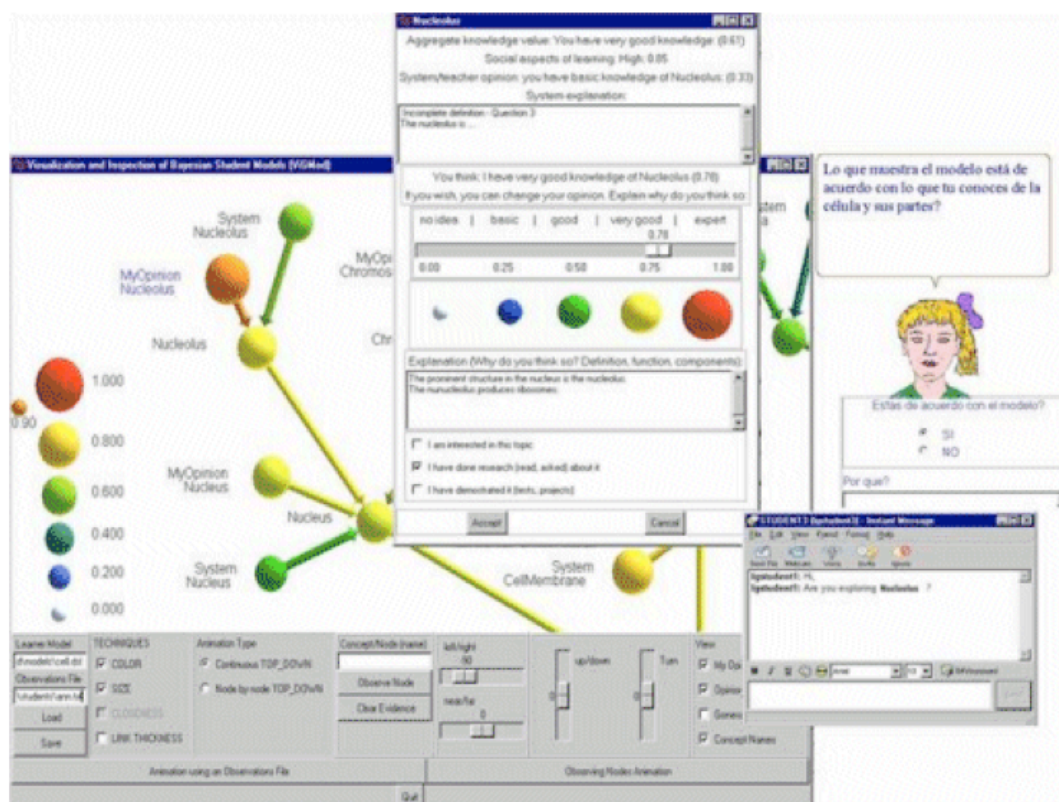


Figure I.10: Use of Bayesian networks and EPAs in OLM, in (Zapata & Greer, 2004)

The use of chatterbots to help the negotiation leading to a collaboratively built learner model between the user and the system's beliefs (Kerly and Bull, 2006), proved useful to visualize the information and enable children to justify themselves in their representation of what they know when disagreeing with the system's assessment. However, due to the difficulties of certain children aged six to eleven with literacy and verbal expression skills (Guha et al, 2004), the use of chatterbots for this PhD's work might not be the ideal use of intelligent pedagogical agents. Could the use of affective cues inserted in the design of traditional EPAs help initiate a communication with the user, based on a few written instructions and a system of proof (performing a test, etc...) entirely graphical, help children this age use better OLM software, and build a learner model more accurately?

Whilst the potential of animated agents to make the experience more engaging is well established (André, et al., 1998; Lester, et al., 1997; Takeuchi & Naito, 1995; Walker et al, 1994) there is also a risk of the user being distracted by it with bad effects on learning (Moreno, et al., 2001; Rickenberg & Reeves, 2000).

Results from (Hall et al, 2007) illustrated the potential of use of emotional interpretations of characters in synthetic characters, and their potential for exploring personal, social, and emotional issues. Research on EPAs has been increasingly interested in ways to integrate affect and affective

support into ITS, from an emotionally-intelligent agent that generates emphatic responses during problem solving (Zakharov et al, 2008; Woolf et al, 2010a) to mirroring student's affective states through non-verbal gestures (Burleson and Picard, 2007). Affective agents are generally preferred to non-affective ones by ITS users (Craig et al, 2002). The inclusion of affect in user-interfaces such as the simulation of emotional expressions by computer agents has been shown to create an emotional interaction between the user and its agents (Brave et al, 2005; Craig, et al., 2002). Pedagogical agents have been used to produce emotional or motivational feedback (Woolf et al, 2009; Woolf et al, 2010a). Students using such agents were reported to increase their math value, self-concept and mastery orientation, with females reporting more confidence and less frustration (Woolf et al, 2010a; Arroyo et al, 2010). Low-achieving students were seen to have greater affective needs than their higher-achieving peers, and improved their affective outcomes with a greater reduction of frustration and anxiety reported when interacting with such agents than the high-achievers participating in the study (Woolf et al, 2010b).

Can interface personas simulating synthetic emotions help children better associate with the agents than non-emotive ones? Can emotive interface personas help children learn better or increase the interaction with the pedagogical tools available? To what extent can the inclusion of affect in the design of pedagogical agents facilitate the learning process and give a more enjoyable learning experience without children being distracted by the task at hand?

1.5 HCI FOR TECHNOLOGY-ENHANCED LEARNING

Interactive products for children regroup three main categories: education, entertainment and enabling products such as web browsers or word processors for children. Special approaches are needed throughout the lifecycle when the intended users are children, which lead to the research in methods for designing and evaluating such products with and for children.

In this section, the issues of interaction and effectiveness of products for children, which are the main concern of the field of Child-Computer Interaction (CCI), will be discussed. Section 1.5.2 will then concentrate on how to create educational software for and with children, prior to report the different methods used to evaluate children products in section 1.5.3.

1.5.1 FROM HUMAN-COMPUTER INTERACTION TO CHILD-COMPUTER INTERACTION

Over recent years there has been a significant increase in the published work relating to children and interaction design. CCI is emerging as a vibrant sub-field of Human Computer Interaction, following pioneering work by Druin and Soloven (1996), Kafai (1990), Papert (1980), Read (Markopoulos et al, 2008), and Rogers and Scaife (Roger et al, 2002). When designing for children, general HCI interface design rules, guidelines and principles can be applied (Dix et al, 2004; Preece et al, 1998). Nevertheless some features of children that differ from adults need to be taken into account when designing software for them (Shneiderman, 2004).

Gilutz and Nielsen (2002) reported for example a large number of cases in which their usability findings for 8-12 years old children highly differed from results generally obtained with adults. Some adaptations have therefore to be made like the words used in heuristics (MacFarlane @ Pasiali, 2005) which led to the creation of new guidelines (Chiasson & Gutwin, 2005a; Druin 1999), principles (chiasson & Gutwin, 2005b), design and usability heuristics (Hanna et al, 1997; Nielson et al, 2005). The problem in using those evaluation techniques is that children will not necessarily fully understand the implications of problems in the interface, due to their lack of experience with the underlying concepts (MacFarlane & Pasiali, 2005).

In the case of a tutoring system for children, it is vitally important that the mode of interaction between the child and computer is efficient and intuitive so that children can access the functionality of the system and focus on the task. It is important that the issues relating to the navigation and understanding of the interface components are reduced, so as to restrain the intervention of the teachers on instructions relating to the task at hand, mirroring the class practices.

Preece et al (2002) claim that "...aesthetics of an interface can have a positive effect on people's perception of the system's usability" and thus a large part of design has to concern the interface in the system.

The interaction between the user and the computer can be seen as a dialogue of communication. It is important that the interaction style is appropriate for the task domain and dialogue type, so that the communication is intuitive and allows the user to access the real functionality of the system.

Druin (1999) concludes that children aged five to seven want interfaces that they can easily control; they want interfaces that 'respect' them, those being interfaces that are not too simple. When designing for children, the interaction style must be suitable for children's ability to read and write using software. With the use of appropriately labelled buttons and symbols, point-and-click interfaces allow easy navigation through the system, as the user should always be aware of "where they are going".

Designers of children's technology must concentrate not only on the mechanics of their interfaces but also on features that will keep children engaged (Chiasson & Gutwin, 2005). Icons can be used to convey meaning as they can present a pictorial clue of the action performed by clicking on them. This allows children to rely on their visual sense of perception and thus make inferences about the action that an icon represents. Work by (Hanna et al, 1997) suggests that the difficulties that children have with reading can be reduced by the use of 'visually meaningful icons', 'thoughtful cursor design', and the addition of features such as rollover, audio, animation, and highlighting. There is some specific research on the selection of icons for children's interfaces. Uden and Dix (2000) and Baecker et al (1991) report that children prefer animated icons, and it is often the case that such icons can offer more information than those with static representations. Some HCI methods designed to include stakeholders into the design process have been adapted for use with children, and are described in the next section.

With the wide distribution of games and other applications on the Internet, as well as the deployment of technology-enhanced applications for education into primary schools, children the world over are becoming more familiar with using computers and computer-based applications from an early age. Researchers in Child-Computer Interaction, and Interaction Design for Children, have investigated how children use technology in order for designers to produce software suitable for efficient use by children. In this context, studies have been investigating different interaction techniques used in children's software. Whilst children appear to adapt to whatever interaction style (IS) is used, some issues remain as to their level of precision in pointing tasks, and their ability to use the mouse in order to fulfil the tasks required by a specific interactive system (Hourcade et al, 2004; Inkpen, 2001).

Interaction is a fundamental part of a computer-based application. Understanding which interaction style improves usability for children is of great research interest. While a wide range of other input devices (pen-based, touch-based, etc...) are investigated and included into current learning technologies, the standard tools present in classrooms and at home remain mouse-based input devices (Hourcade et al, 2007). Therefore, this research agenda and studies concentrate on using direct manipulation techniques, with mouse-based input devices.

In the 1970s and early 80s with non computer-based studies undertaken by psychologists (Kerr, 1975), and then in the 90s with computer input device investigations (Hourcade et al, 2004), studies clearly showed that pointing performance increases with age during childhood. These results motivated researchers in defining interaction methods specifically designed for children's developing skills. The aim of numerous studies has been how children appropriate various types of interaction styles within direct manipulation, with children from different backgrounds, experience in using computers prior to the studies, and age groups. A few are outlined here: Berkovitz (1994) illustrated the difficulties for children aged six to twelve years old to manipulate "Marquee Selection" techniques. Issues relating to accuracy of pointing tasks for younger children (Read et al, 2002), the size of the mouse (Hourcade et al, 2004), or maintaining pressure on the mouse-button for a prolonged amount of time have been investigated. Inkpen's study on the drag-and-drop versus point-and-click techniques with children aged nine to thirteen years also showed accuracy problems in dragging and pointing tasks (Inkpen, 2001).

In order to maximize software usability and its integration into school practices, information gathering is paramount at different moments of the design process, and leads to considering the benefits of involving children and teachers in the conception of the applications.

1.5.2 CHILDREN AS DESIGN PARTNERS

Designers in the CCI field advocate that involving children in different phases of the design process is vital to ensure that most of the requirements needed are included (Mazzone & Read, 2005).

A user-centred design assumes that the emphasis during the lifecycle is on engaging with the potential users of the product. This engagement can take place at the early requirements and analysis stages and at the point where the final product is user-tested, but, it is more crucial during the design process results in a design being proposed, evaluated, redesigned, re-evaluated, and so on, until a design is reached that satisfies the team's criteria well enough to be implemented.

This process is called iterative design, as illustrated in figure I.11.

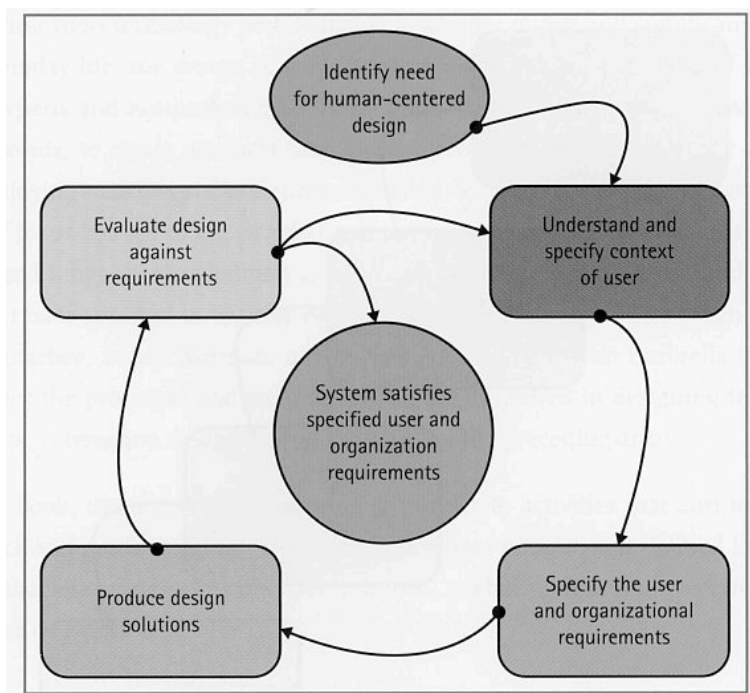


Figure I.11: User-centred iterative design

Research has been carried out in this area to explore the suitability of design techniques for involving children in specific phases of design.

Druin (1999) introduced a model of the roles of children in the design process, illustrated in Figure I.12:

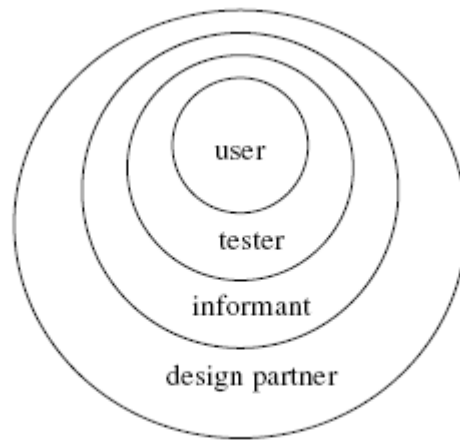


Figure I.12: Roles children can play in the design process

The four categories of children's involvement in the design process are defined in (Mazzone & Read, 2005) as:

- *User*: Users are observed performing existing activities and/or in specific or natural settings while they use the prototypes. This can take place at the beginning, during, and on completion of the design process, through user observation and ethnography methods.
- *Tester*: Children are observed testing technologies and asked to provide feedback through interviews, questionnaire and think aloud techniques, most commonly used at the end of each development phases.
- *Informant*: Children are seen as experts, informing designers of key issues related to their experience, helping to develop early design ideas and testing prototypes in development.
- *Design partner*: Children work as equal members of the design team, helping in identifying problems and solutions to improve the technology they may need in support of their activities.

Cooperative Inquiry (Druin 1999) takes elements from both Contextual Inquiry and from Participatory Design and adapts them for use with children. It investigates changing the way children are consulted regarding the development of software that affects them. Druin (1999) found that children in the 'concrete operational stage' to be the group most effective when prototyping, as they are "verbal and self-reflective enough to discuss what they are thinking...(not) too heavily burdened with pre-conceived notions of the way things 'are supposed to be'." According to Druin (1999), it is important that adults and children work together on the design, as no one partner should be more important than another. However it is important to bear in mind that children are not "little adults" and that care must be taken when planning design sessions to ensure that the activity is set at the right level to enable the children to be able to express their ideas in a useful format. Using participatory design with children has proven to be a useful technique along with cooperative inquiry and helps to design systems more effective and appealing for the future users. Methods such as PICTIVE (Muller, 1991) or C.A.R.D.S. (Tudor et al, 1993) proved to be useful for use with children and adults with an appropriate choice of level of desired requirements specifications and session outcomes. One point to note is that the design materials need to be carefully chosen to match the activity, as one standard set of equipment is unlikely to be suitable for all ideas or participants. The use of C.A.R.D.S. to examine the system workflow might not be appropriate for younger children with no experience of interactive systems.

Most participatory-design methods aimed at child design partners use, insist on the use of different techniques and adaptations to the child learners. In KidReporter (Bekker et al, 2002), the aim of the described design process is to create a method:

- *Motivating and stimulating for children:* the children should enjoy the whole process and get something out of it
- *Suitable for children's reading, writing, and verbal expression level:* this method will need to take into account the level of expression of the youngest children
- *Optimizing the quality of data:* When working with children to gather some usability or data on a specific subject, it is essential to check for the consistency of answers, taken from different sources such as interviews, free expression, or observations.

1.5.3 EVALUATING SYSTEMS WITH AND FOR CHILDREN: THE IMPORTANCE OF FUN

The methods used to evaluate a product are determined by the stage of development of the product at the point of evaluation, the availability of resources (including users and design experts) and the feature(s) of the product that are being evaluated and the purpose of the evaluation to be performed.

Evaluator methods in the CCI field contain HCI general methods adapted to children (Heuristic Evaluation, Cognitive Walkthroughs, Structured Walkthroughs, Think Aloud, Surveys) and some methods especially designed for children (Co-discovery, Peer-tutoring). They can be separated into two groups, evaluating either by observing or asking children.

In observational work with children, the age, the position, and the appearance of the observer may influence the findings (Hanna, et al., 1997). Some general characteristics of children that need to be taken into account when carrying out evaluations are that they have developing capacity to verbalise, they have different levels of extroversion, their knowledge and skills may be different, and although generally being very honest in their judgements, the reliability of reported data is questionable as they are closely linked to their developing personality and show numerous examples of personal differences rather than group tendencies (Druin, 1999; Markopoulos & Bekker, 2003).

Rode et al (2003) highlight that any usability test on pedagogical software to be performed in classroom settings must fit within the school sessions; that it has to fit the framework of the school and that it should not cause ‘lesson stoppers’. Rode et al(2003) use this term to describe events or materials like games, animations, technology failures, etc., that distract children from curriculum objectives and could derail a lesson.

Hanna et al’s (1997) findings on the potential of using children as usability testers revealed that children from the “Elementary School” stage (aged six to ten yrs) are relatively easy to include in software usability testing. Their experience in school makes them ready to sit at a task and follow directions from an adult, and they are generally not self-conscious about being observed as they play on the computer. They will answer questions and try new things with ease. In this age range, children will develop more sophistication about how they can describe the things they see and do. Six- and seven-year-old children will be more hands-on, i.e. ready to work on the computer but a little shy or inarticulate when talking about the computer. Ten-year-old children may have extensive computer experience and be ready to critique your software.

Children have their own abilities, curiosities, needs, skills and expectations. Their goals while using computer are education or entertainment rather than a common adult goal of productivity (Chiasson & Gutwin, 2005a). New criteria have to be taken into account when evaluating products for them such as fun. Children from the concrete operational stage are considered old enough to use relatively sophisticated software, but still young enough to appreciate a playful approach (Schneider, 1996).

Shneiderman (2004) states that “Fun-filled experiences are playful and liberating – they make you smile. They are a break from the ordinary and bring satisfying feelings of pleasure for body and mind”. Including a notion of “fun” into the design process can increase the motivation of the child to use the software, which leads to the inclusion of fun features in the design such as “alluring metaphors, compelling content, attractive graphics, appealing animations and satisfying sounds”.

Many researchers have explored the relationship between fun, play and learning, reasoning that fun contributes to being motivated to pursue an activity, and as such can also contribute to learning effectively (Malone & Lepper, 1987). Draper (1999) suggests that fun is associated with playing for pleasure, and that the sense of fun may be achieved by presenting tests in a different medium incorporating multimedia stimuli.

Developing educational software incorporating a gaming genre is seen as a motivational factor for children enticing them to use the software (Alessi & Trollip, 2001). The intervention of pedagogical agents at appropriate moments in the software might motivate the children and help them relate more to the software storyline. It might encourage them to use software not only as a tool to help their learning but also as a game. The use of EPAs creating an emotional attachment from the children could help them view the activities more like a game, and increase their level of fun and motivation to play.

1.6 SUMMARY

In this chapter, background research concerning the following areas of the problem domain is discussed: theories on learning and teaching through the use of technology, development and use of Open-Learner Modelling techniques the use of affect for learning; and how to build and evaluate technological applications for and with children. From the study of current research on all aspects of the problem domain, a number of research questions have been drawn, to be answered in this PhD thesis:

RQ1: What theories of emotion can be followed to produce a model of EPA's affective response in OLM software suitable for mathematical drill-and-practice applications?

RQ2: How can user-centred and participatory design techniques be used with children and their teachers to create affective components to be integrated in tutoring systems for children?

RQ3: What is an appropriate representation of a learner-model (LM) for a child to interact with?

RQ4: How can user-centred and participatory design techniques be used on children and their teachers to represent the Learning model components of OLM tutoring systems for children?

RQ5: Does opening the content of the learner model to the child-user facilitate learning and motivation?

RQ6: How can the use of pedagogical agents or the inclusion of affect in the design of OLM tutoring systems help children interact with their LM, facilitate learning and motivation?

RQ7: How can children interact with their learner-model?

RQ8: What level of inspection and modification of the learner model content by the child-users is suitable for the child to learn efficiently?

RQ9: What kind of learner model inspection/modification by users is most suitable to help children make use of the learner model to learn more efficiently and gain more reflection on the pedagogical activity: full editing (editable learner model), no editing (inspectable learner model), or building the model with a user-system negotiation process (negotiated learner model)?

The work presented in this thesis is scoped to children users aged seven to eleven years old, drill-and-practice mathematical OLM applications deployed directly in classrooms, and children's view and use of the learner-model with inspection and/or edition rights. These questions will be addressed by a combination of theoretical analysis, system design and empirical evaluation.

The empirical studies will follow Barnard & Sandberg's (1996) actions to overcome the lack of self-assessment found in learners interacting with OLM:

- *Provide learners with the means to undertake self-assessment*

In chapter 5, a study will investigate what children make of the information offered to them in an inspectable learner model application.

- *Investigate how learners make use of such means*

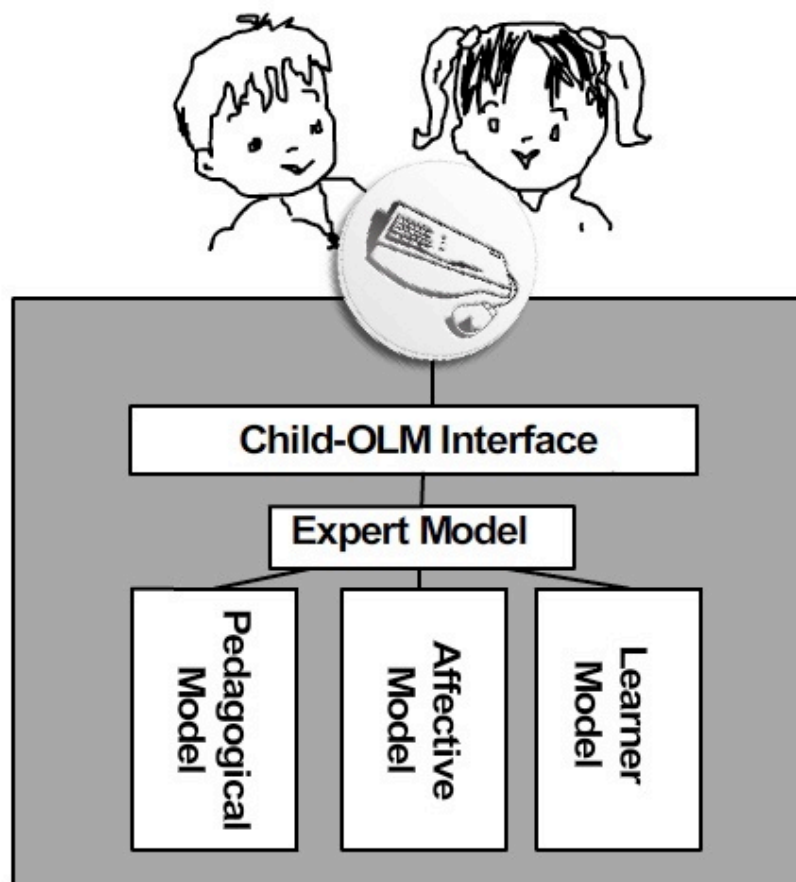
In chapter 6, different interaction mediums will be tested between the user and the learner model, in an inspectable learner model application.

- *Investigate how learners can be stimulated to use the facilities offered in a fruitful way*

In the last empirical chapter, chapter 7, three types of OLM (inspectable, editable, and negotiated) applications, and their use for better learning, increase motivation, and reflexion on the learning achievements and learning process will be investigated with children aged 7 to 11.

Chapter 2 will now present the two OLM applications later used in the empirical studies investigating the aforementioned research question.

Chapter 2: Two OLM platforms: *DividingQuest* and *Multipliotest*



2.1 INTRODUCTION

This PhD is concerned with the potential use of affective OLM pedagogical environments by young users (6 to 11 year olds) to learn and practice mathematical skills. In order to study how children use such systems, two OLM intelligent tutoring systems have been created/updated in the course of this PhD: the *DividingQuest* and *Multipliotest*. Both were designed in participatory-design with teachers and children initially for an MSc dissertation (Girard, 2006), the applications were then further developed to accommodate the needs of this research work in terms of the experimental investigations.

In this chapter, all aspects of the applications are presented, in terms of software architecture (section 2.2), elements of the application relating to the Child-Computer interaction and interface (section 2.3), concluding in section 2.4 with a succinct description of all the components of the architecture: pedagogical, affective, learner model, and user-interface components.

2.2 SOFTWARE ARCHITECTURE

Software architecture for the *DividingQuest* and *Multipliotest* was defined following Barnard et al's (1990) methodology for open learning environments in which the learner is located at the centre of the learning environment. The description of the environment corresponds to Van den Brande's (1993) view that open learning involves accessibility (in terms of time and place), learner centred pedagogy, and learner's control of context.

The architecture enables all the components of the OLM created to adapt dynamically to the changing student model. A series of issues were reported by Brusilovsky (1995) on using a single representation of student's knowledge within the architecture in terms of software maintenance, and evolution of the learner model. His solution consisted of a student model centred architecture, based on several LM projections, made some progress in answering this matter, but produced an architecture growing in links between components, redundancy of views, and lack of genericity for student's changes in terms of preferences or learner model content.

Figure II.1 illustrates the architecture for affective OLM environments proposed in this PhD work, which includes 4 main components - the pedagogical model, affective model, learner model, and user interface – as well as an expert system linking everything together.

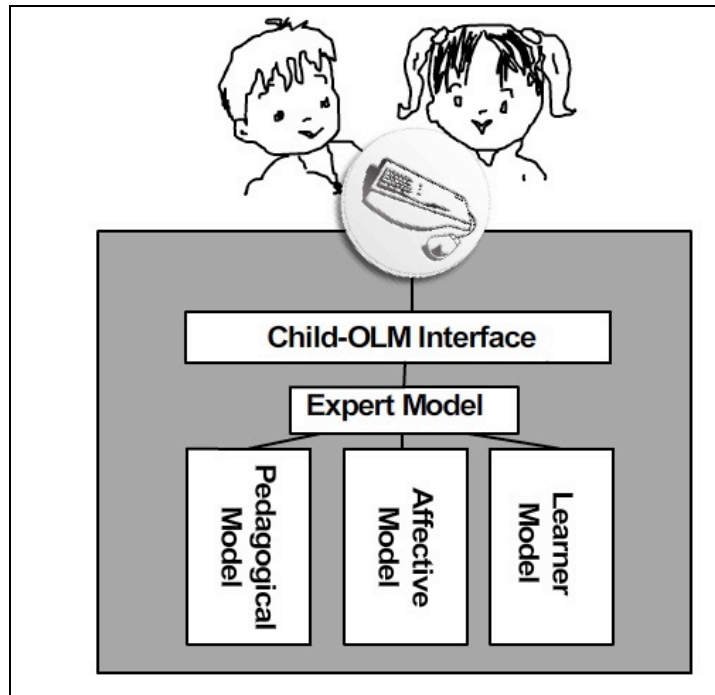


Figure II.1: OLM architecture for *DividingQuest* and *Multipliotest*.

A multitude of information is needed in each pieces of software in the component usually defined as the “expert system”: not only does it include pedagogical content and future user instructions/feedback, but also affective feedback, as well as records of the LM child-interactions. Furthermore, the experiments to be performed in the empirical studies needed the architecture to enable several methods of input in the learner model while recording information in its whole. For this reason, the previously cited approach to Open-Learner Modelling system architecture was coupled with a common Human-Computer Interaction system architecture: the Model-View Controller – MVC – (Krasner & Pope, 1988; Dix et al, 2004):

- The *model* represents the application semantics;
- The *view* manages the graphical and/or textual output of the application;
- The *controller* manages the input.

Figure II.2 details the architecture in terms of using the MVC model: When the user interacts with the OLM, the main *controller* agent receives all information concerning the context of the action from the agents linked to the user-interface. It includes the following information: type of action (e.g. begin an activity, quit a game, ask for help, visualize the LM, etc...), information transmitted from the current state of information (e.g. the answers given once validated by the user, in order to produce feedback on this answer and go to the next question), etc... It then transmits the data to the *model* by category of information (e.g. affective, pedagogical,...), and signals to the *view* the next action to be

done. The *model* part then updates each component of the application, as well as the information it holds about context of interaction. By request of the *controller*, information about the current context of interaction is taken from the *model* and transmitted to the *view* in order to build the interface for the next action, as detailed in the following example:

Figure II.2 illustrates the architecture chosen, with a detailed description of the MVC component:

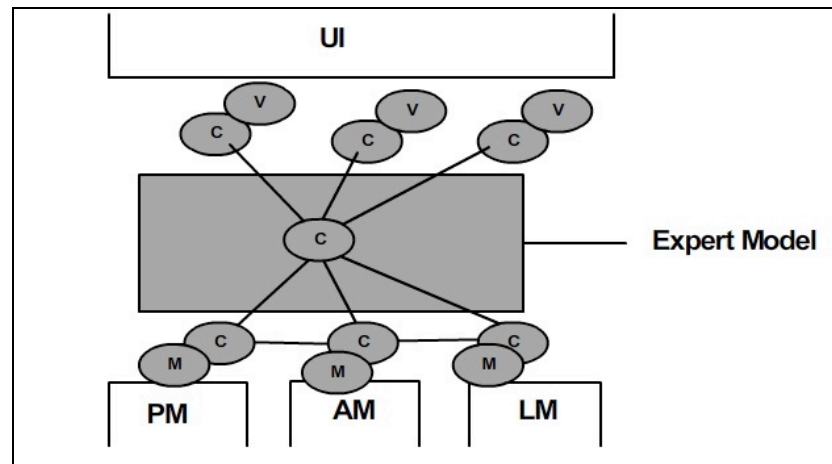


Figure II.2: MVC applied to OLMs.

Here is a simplified example of the MVC process when the child just answered a question in one of the activities:

- The main agent's *controller* registers the context from different *Controller-View* agents linked to the interface: the activity and the question as well as the answer, the experimental condition the user is working on, and the user's identification.
- The *controller* then initiates a dialogue with each *Model-Controller* agents of each sub-component of the architecture:
 - The agent linked to the pedagogical model calculates the answer of the question and defines the pedagogical feedback to be given according to the user's aptitudes at answering the question given in input, and the type of feedback this user follows, defined in the learner model component, values kept in the main controller as current context.
 - A *Model-Controller* agent linked to the learner model is then called to update the child's accomplishments concerning the action and information about the

child's past actions (feedback given, interaction..) is transferred to a third *Model-Controller* agent connected to the affective component.

- Using the information about the user's characteristics, feedback level proposed, and feedback the child received in the past, an affective feedback is selected and transmitted to the main *Controller*.
- The *controller*, after updating the new information about user context, then communicates pedagogical and affective feedback, as well as the next action to be undertaken on the interface, to the *Controller-View* agents to modify the different parts of the user interface and give the user the answer, as well as the feedback corresponding to what s/he previously inputted.

Once the architectural structure was decided upon for the OLM applications manipulated in this PHD, issues related to the Human-Computer Interface and Interaction, summarized in the next section, were investigated. The aim was to define the main graphical and interactive characteristics to be shared by the different applications, in order to reduce any possible bias in the data gathered in terms of software usability and coherence.

2.3 HCI-RELATED ISSUES

OLM Intelligent Tutoring Systems are complex educational applications. When conducting an evaluation of such systems, multiple factors can be found that impact the results. They are linked to the three dimensions of evaluation: usability, utility and acceptability (Tricot, 2003). In order to reduce the impact of different user-interface designs on the results on the results of the various experiments reported in this PhD, all the software interface used should be as close a possible in terms of pedagogical, HCI, and OLM considerations. The design choices made when building the various interfaces of the two OLM educational software (*DividingQuest* and *Multipliotest*) are summarized.

2.3.1 DESIGNING SOFTWARE FOR FRENCH AND ENGLISH USERS

Since the Shengen treaty (Shengen, 1985), enabling people from European countries to travel freely from one country to another (within the Shengen space), French school population has been diversified in terms of children's nationality and cultural background. Similarly, with the development of low-cost airplanes between France and the United Kingdom, and UK citizens choosing to retire or settle in French country-side towns, a lot of English families have settled in the rural areas. Their children are divided between English and French schooling systems and brought cultural changes in the French classes. Any pedagogical software deployed in French schools would therefore need to take into account the pluri-cultural background of its students.

DividingQuest and *Multipliotest* are web-based computer-assisted drill-and-practice learning environments with an open learner model to support the learning of a specific domain of mathematics (divisions and multiplications) for primary school children (aged 9-11 and 7-11). The first software, *DividingQuest*, enables the study of divisions concepts, and is employed in chapter 6 with English children aged 10-11 years old. The second software, *Multipliotest*, used in the evaluations of chapters 5 and 7, is employed on French children aged respectively 9 to 11, and 7 to 9 years old, who worked on multiplications.

The choice of using first English and then French children lies in the way the applications were developed, and the possibility of comparing children's reactions to OLM software according to their national school system.

2.3.2 DESIGN OF THE APPLICATIONS

Both pieces of software were initially designed using participatory-design techniques with both teachers and children from English (*DividingQuest*) and French (*Multipliotest*) classes. This satisfied a certain conformity with the pedagogical methods used to learn and practice the concepts, included within the applications, as well as the creation of software usable for both children and their teachers. This aims at facilitating their deployment to and effective use in classroom settings (Van Dam et al, 2005). It also helped produce software determined by the children's interests to increase their motivation in using such applications.

The participatory design sessions included teachers and children as design partners, informants, and usability testers. The design process of some of the components, inspired from research on participatory-design using children in the design process (Druin, 1999, Kelly et al, 2006, Bekker et al, 2002) will be discussed in context in chapters 2 through 4, and in the Appendixes. A

discussion on field results concerning using participatory-design techniques with children and teachers for educational purposes is presented later in the conclusion chapter.

2.3.3 TRAFFIC-LIGHT SYSTEMS AND SMILEY FACES: DEFINITION OF THE METAPHORS

When developing a tutoring system, a crucial decision lies in the choice of learning metaphor(s). Metaphors help students associate interface elements with what they should do, what they have learned, and with concepts not yet fully understood. Metaphors are especially important in product design and user-interface design in order to convey the right message as to how to interact with the tool, and in the user's definition of the goal of each functionality considered (Shneiderman, 1998). For this reason, when thinking about designing affective products and tools, it is important to choose one (or several) metaphor(s) to increase the understanding of the information presented in software.

The two metaphors of learning most widely used in schools in England are the Traffic-light system and the Smiley Faces, illustrated in Figure II.3 with their meaning in terms of learning. They are used to represent children's performance while undertaking a task or grasping a concept, dividing the development of learning for a specific concept into three states, also known as pedagogical targets:

- the concept is not understood; 😞
- whilst the student is still making errors, s/he begins to partially understand; 😐
- the concept is grasped. 😊

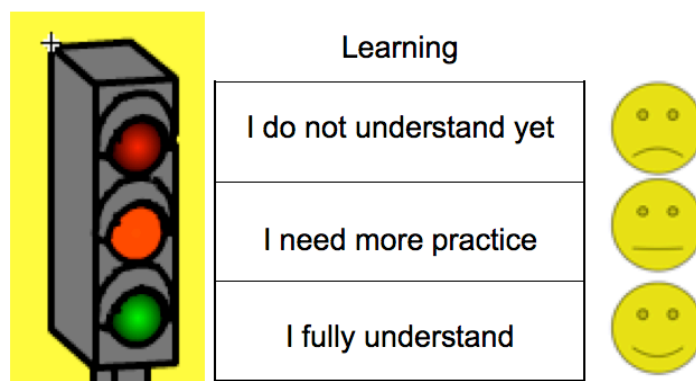


Figure II.3: The three states of the *Traffic-Light System* and the *Smiley Faces*

The Traffic-Light System, relies on the international symbol followed for road safety, and associates a colour (red, orange, or green) to a learning state (not understood, partially understood,

understood). The second metaphor, the Smiley Faces, is associated with an emotion (sad, neutral, happy).

Children's understanding of the two metaphors was tested empirically (cf. Appendix B). The results can be summarized as follow:

- French and English children aged seven to eleven have a good understanding of each metaphor separately
- When confronted to a combination of both metaphors included in the design of educational components, all French and a majority of English children aged seven to eleven participating in the study could correctly associate graphical icons (designed with both colours and facial expressions) with their corresponding pedagogical targets.

Both metaphors will therefore be used to represent pedagogical feedback in software and highlight learner model information presented. The use of either metaphor in software for specific aspects will be given rationale in the experiment chapters, and will differ according to the research goals investigated.

2.3.4 EMBODIED PEDAGOGICAL AGENTS

In this PhD, the potential of using Embodied Pedagogical Agents (EPAs) as an interaction medium with the learner model will be investigated, for young users aged seven to eleven. Research has shown that EPA's "look and feel" (e.g. visual representation, communicative style, and underlying pedagogical goals) impact how users perceive and use the agents, as well as learners' self-confidence and transfer of learning (Baylor, 2005; Baylor & Kim, 2005; Gulz & Haake, 2006; Isbister, 2006; McCloud, 1993). Therefore, when defining an EPA for use in specific educational software, the visual and pedagogical aspects of the agent's design will have to take into account the end-users performances and characteristics.



The study presented in Appendix C investigated users' preference as EPA for children aged seven to eleven. The results can be summarized as follow:

- Children aged seven to eleven will favour more naturalistic (humanoid-shaped) instructors, and stylized (smiley-shaped) learning companions, both EPAs with a low level of detail in the visual static appearance.

- When considering the EPA's pedagogical role, a preference can be seen for learning companions as opposed to a role of instructor/teacher. Children indeed expressed a greater level of trust and forming a closer bond to the characters.

The design choices concerning all EPAs used in the *DividingQuest* and *Multipliotest* were taken from the results of the literature and the previous study, illustrated in table II.1. The table content is detailed and rationalized in the Appendix C.

Table II.1: Description of the EPA's design

	<i>DividingQuest</i>	<i>Multipliotest</i>
Visual Static Appearance		
Basic Model	 <p>Séléna magical humanoïde creature (witch)</p>	 <p>Multiplio multiplication symbol</p>
Physical Properties	Age: Immortal but looking 15-20; Caucasian ethnicity; clothes, body, and hairstyle related aspects corresponding to the software fantasy theme and used for the inclusion of colours.	No age; presence of a face with eyes, eyebrows, nose and mouth onto the multiplication symbol, which are used in form and colour to represent affect.
Graphical Style	<ul style="list-style-type: none"> • dynamic representation; • simplified and low level of details; • stylized and non-realistic; 	
Pedagogical role		
<ul style="list-style-type: none"> • Expert for navigation and learner-model representation; • Mentor when giving user feedback on his/her progress and facilitating help-seeking behaviour; • Guide and motivator when considering pedagogical choices in interaction; <p>In overall, moderate authoritative role, learning companion</p>		
Communicative style		
Pedagogue mainly task-oriented, but trying to create an affective relation with the user: low level of argumentativeness; high degree of assertiveness; more descriptive than narrative; extrovert; and quite connected to the user.		

In the first version of the *DividingQuest* application, the EPA, the witch 'Séléna', was integrated in the environment, and tested in usability studies with English children aged nine to eleven (Girard, 2006). The character was perceived as an authority figure, assimilated to the teachers the children obey instructions from. They represented a guide holding the knowledge when perusing the activities, and able and willing to help the users along the way, navigation-wise. However, while the colours were well perceived and understood as the emotional valence of the witch's feelings, the facial expressions were found hard to read. For this reason, and to conform to children's preferences in humanoid-shaped instructors with a low level of details, the following modifications were made in this PhD's version of the EPA: the head was made bigger, in order to enable better visualization of the facial elements that form the facial expression.

The A.M.E.R. model (introduced in section 2.4.2) defining the different emotional states to be expressed by the EPA, was then created. To follow, a participatory design process, introduced and detailed in section 3.3, involving children from a French primary school (children from French and English nationality, aged seven to eleven years) was performed in order to redesign the graphical animations using the witch ‘Séléna’ to portray emotional feedback in the *DividingQuest*.

For the *Multipliotest* application, a study involving twenty-nine children (from a French primary school, aged seven to nine) as design informants was performed in order to create an EPA for the application. The design session, presented in details in the Appendix C, produced ‘Multiplio’, a smiley-shaped learning companion with a low level of detail in the visual appearance, as illustrated in the right side of table II.1.

2.3.5 INTERACTION STYLE AND USABILITY

A critical Human-Computer Interaction decision concerning the OLM applications is the choice of interaction style, as some styles are known as potentially problematic for young users. Whilst children appear to adapt to whatever interaction style is used, some issues remain as to their level of precision in pointing tasks, and their ability to use the mouse in order to fulfil the tasks required by a specific IS (Hourcade et al, 2004; Inkpen, 2001). The targeted children age-groups for this PhD are children between seven and eleven years of age. Classified according to Piaget (1974) as belonging to the 'concrete operational state', they are considered able to use simple keyboard interaction, and, as their ability to learn to type grow throughout this age group, a relatively fine control of the mouse (Schneider, 1996).

This PhD's work falls in the context of “in-class” field studies, and in order to produce software easily deployable in classroom, and reusable by teachers outside from experimental settings. For this reason, “traditional” desktop computers, with mice and keyboards, were selected. When considering direct manipulation interaction styles with a mouse and keyboard in pedagogical settings, the following styles are usually chosen:

- point-and-click (the answer is on the screen, the user positions the mouse on the target and clicks),
- drag-and-drop (the user selects an answer by clicking on the mouse, and without releasing mouse pressure, 'drags' the item on the interface in the answering area, before 'dropping' the item by releasing the mouse button),

- point-and-select (the answer is to be chosen between a number of possible answers, with only one visible at a time. A scrolling-tab, moving over the chosen answer, and selects it by clicking on the mouse), and
- keyboard-typing (writing an answer using the keyboard).

Due to the fact that a subset of the children who participated in the studies are new to reading and writing. A design choice to avoid children “writing an answer” or reading too large or too difficult an amount of text was made. All instructions and questions feedback either used sound data duplication (for the *DividingQuest*) or relied on the use of visual cues to aid children’s understanding.

Inkpen’s study (2001) on the comparison between the drag-and-drop and point-and-click techniques showed that children aged nine to thirteen showed accuracy problems in dragging and pointing tasks, performing better with the point-and-click interaction style. The use of the drag-and-drop technique has therefore been avoided in this PhD’s work.

For the choice of the application’s interaction style a study, presented in details in Appendix D, was performed in order to investigate the two remaining interaction styles considered: point-and-click and point-and-select. The results are summarized as follow: The point-and-click interaction style is a more effective mouse interaction style than the point-and-select style for children aged seven to nine years, in terms of achievement, number of interaction errors, speed and accuracy of answers, as well as children’s preference.

A single interaction style was therefore implemented in software: point-and-click interaction. From an HCI point of view, all features were evaluated in usability against heuristics, design guidelines, and others from the literature in terms of choices of icons, sizes and place of elements, use of colours, and so on... (Shneiderman, 1998, 2004; Nielson et al, 2005; MacFarlane and Pasiali, 2005; Dix et al, 2004). All pieces of software were evaluated in terms of usability by teachers, children, and HCI experts, using think-aloud protocols.

The decisions aforementioned underpinned the design of two pieces of OLM software with reasonably close user interface components and pedagogical strategies. The involvement of users from the end-users age group and nationality, as well as their teachers, in the whole design process should help in the design of educational software more usable and enjoyable to use (Markopoulos et al, 2008).

2.4 AFFECTIVE COMPUTING IN OLM APPLICATIONS

In this PhD, the investigations relating to affect are two-fold as regards to the affective component of the OLM applications, presented in this section:

- The creation and use of an affective component that enables the capture of what the users are feeling when interacting with software: *Sorémo*.
- The definition and use of an affective model to guide the feedback given to children when interacting with software for better learning and motivation: the A.M.E.R. model.

This section presents the two aspects of affect in the applications, their design, evaluation, and rationale. In the first part of the section, the non-verbal self-report tool developed during the PhD to capture children's emotional states while learning on OLM software is introduced.

2.4.1 *SORÉMO*, EVALUATING CHILDREN'S EMOTIONAL STATE

When using any educational product, children experience emotional situations and feel emotions of different arousal, and valence (Woolf et al, 2010b). Methods such as the Self-Assessment Manikin (Sander & Scherer, 2009), or the AffectButton (Broekens & Brinkman, 2009) would enable children to express their current emotions by selecting a single icon or state to represent the affective moment. However, the selection of a unique representation of an emotion as an emotional state might not suit best children's understanding and mental representation of such states. This is particularly salient when considering mixed-feeling emotions (as discussed in the literature review, section 1.3.3). In this section, a new self-reporting tool, *Sorémo*, is presented. It allows the representation of emotional states as a combination of basic emotions, varying in terms of valence, arousal, and intensity of the emotions reported.

A first study, presented in Appendix E, investigated children's and adults' understanding of emotions when embedded in digital characters. The digital characters of the PrEmoTM system (Desmet, 2001) were chosen for this experiment due to the following considerations: as the usability method claims to be used cross-culturally and does not ask users to verbalize their emotions, which facilitates its use by children with low literacy skills. The results can be summarized as follows: While most negative emotions were generally well recognized by all participants, the positive emotions were identified very poorly by all participants, and confused with emotions from the same valence. On average, children seemed to be able to recognize far fewer emotions than the adults. The results seem particularly contradictory between the adult and child populations for three emotions: *unpleasant*

surprise, fascination, and inspiration. It would seem that the visual cues and affective facial/bodily expressions offered by the PrEmo system are not sufficient to help children recognize the 14 emotions represented, especially with regards to the positive emotions.

Designers in the CCI field advocate that involving children in different phases of the design process is vital to ensure that most of the requirements needed are included (Mazzone & Read, 2005). The results of the study described in Appendix E showed that children and adults do not recognize, or understand, affective states at the same level. Therefore, when designing affective educational products, using children from the product's targeted age-group should be beneficial to assure a conformity in the desired and actual product use. However, when considering using affect to represent educational feedback, children alone are not enough to design a product of good pedagogical value. The inclusion of teachers working with children of the appropriate age group should help in the definition of an emotional set corresponding to the pedagogical goals pursued.

SORÉMO, THE INSTRUMENT

Sorémo (in French: “*Sorcières émotives*”, “Emotive Witches”) is a non-verbal, self-report, affective method that measures the user's emotional state according to nine emotions represented by user interface characters (witches). This method has been especially designed for child users of educational software products to investigate the relationship between learning and the emotional states reported. Appendix F presents in details the user-centred iterative process, using children and teachers as design informants, and children as usability testers, that led to the creation of a set of emotions and emotional representations to be included in the new instrument.

THE SORÉMO EMOTIONAL MODEL

The instrument graphically represents nine emotions, classified here along the dimension of ‘pleasantness’: four positive emotions (*happy, captivated, satisfied, inspired*), four negative ones (*puzzled, bored, disappointed, angry*), and one of neutral valence (*thoughtful*).

These emotions occupy different positions on the dimensions ‘pleasantness’ and ‘arousal’ (physical state of engagement) as described by Schlosberg as the two most accepted dimensions of emotion (Schlosberg, 1952). In an early study, Ainley (2005) used an emotion probe based on Izard's (1977) differential emotions theory, with one face icon representing each basic emotion (sad, surprise, interested, scared, embarrassed, angry, happy, bored, disgusted). A “neutral” face icon was added to allow student to report feeling nothing in particular (Izard, 1977). With academic tasks, interest, neutral, and bored were the most commonly reported states (Ainley, 2005). For this reason, an emotion of neutral valence was defined, thoughtful, enabling children to report on this neutral state.

The emotions used in the *Sorémo* instrument can be considered in terms of two dimensions: the level of ‘pleasantness’ and ‘arousal’, as illustrated in Figure II.4.

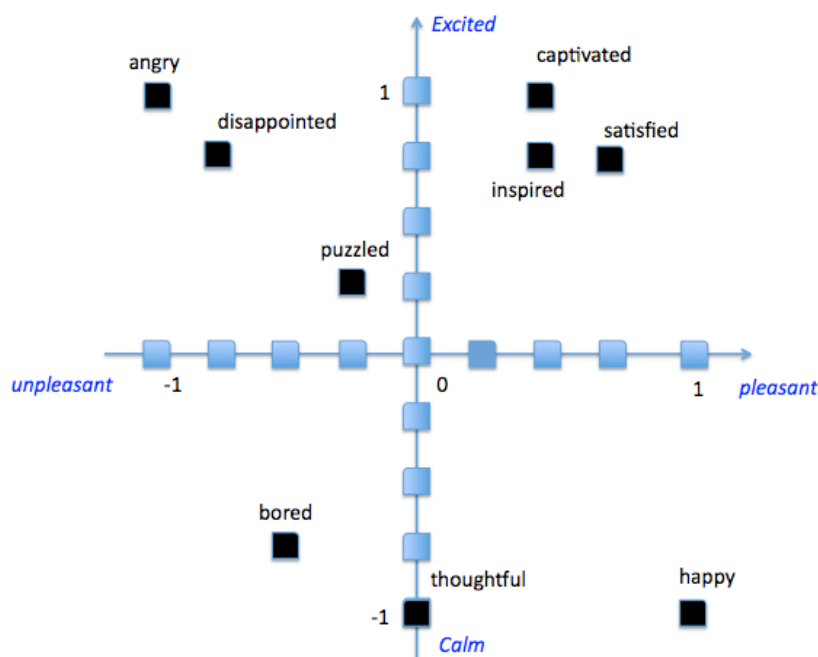


Figure II.4: Representation of *Sorémo* emotional states according to the axis of pleasantness and arousal

Each emotion is designed in the 2D plane according to both levels, following the PAD theory (Mehrabian, 1980; Osgood, 1966). Each emotion can be represented in terms of arousal and pleasantness, and a third dimension is added: the intensity of the emotion.

A pondered score of arousal level, and pleasantness level for one emotional state reported can then be formed, using a middle scale as half-point scale and a “strong” scale as a full point. The resulting number falls between $[-2.5, 2.5]$ for both axis. Table II.2 represents one scaling of a child’s emotional state using *Sorémo*:

Table II.2: *Sorémo* scaling in terms of strength.

	Zero	Middle	Strong
Angry	X		
Bored		X	
Captivated			X
Disappointed	X		
Happy		X	
Inspired			X
Puzzled	X		
Satisfied			X
Thoughtful	X		

The mean levels of pleasantness and arousal, once calculated, can be represented as in Figure II.5:

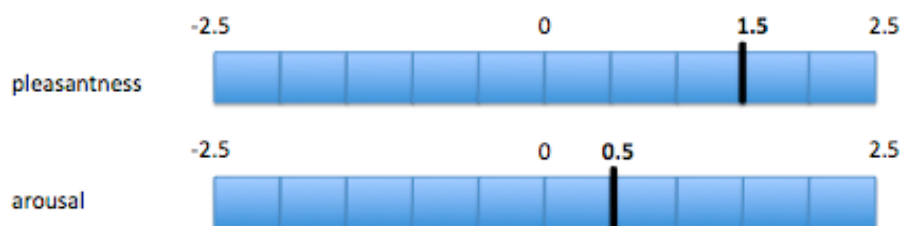


Figure II.5: mean levels of pleasantness and arousal

SCALING SYSTEM, INTEGRATING THE AFFECTIVE COLOURED METAPHOR

Sorémo was designed using two of the most popular metaphors for learning used in primary schools: The Traffic-Light System, and the Smiley Faces, described in section 2.3.2. Each of the nine emotions pictured in *Sorémo* has been designed using dynamic bodily and facial expressions, in addition to the representation of verbal cues. Increasing the emotion representation capacity, and children's ability to understand and recognize such emotional states, colours and emotions were combined to represent the strength of an emotion felt, as described thereafter in the '*Sorémo* procedure'. Indeed, when using *Sorémo*, a child is asked not only to select one or several emotions that correspond to her/his current emotional state, but also to scale how deeply s/he feels this emotion, by selecting a colour/smiley for this emotion. As previously stated, the emotions in *Sorémo* are represented along the two dimensions of 'pleasantness' and 'arousal'. The Traffic-Light metaphor has also been embedded in the design of the instrument to emphasize the valence of the emotions represented and facilitate the classification of the emotions represented:

- On the one hand, the S-shaped in the background has been laid down on a Traffic-Light like shape, each block separating the most positive, negative, and then neutral emotions, as illustrated in Figure II.6. The most negative emotion, *angry*, is placed on the top-right end of the S whereas the most positive one, *happy*, is on the bottom-left end of the shape to follow traffic-lights. The three most negative emotions (*angry*, *disappointed*, *bored*) occupy the red part of the Traffic-Light background, whereas the least negative emotion, *puzzled*, links the negative and the neutral emotions on the orange section of the Traffic Light. The least positive emotion, *inspired*, also on the orange section, links the neutral to the positive emotions, while the most positive emotions are placed on the green section from right (*captivated*) to left (*happy*).
- On the other hand, each pictorial representation emphasizes the valence of the emotions. The metaphor being displayed as feedback, in the form of the colour of the witches' clothes: the colour of the witch's clothes change in function of the valence of the emotion (the more the emotion represented is positive, the more the clothes of the witch are green).



Figure II.6: The instrument interface

A CROSS-CULTURAL INSTRUMENT

The cartoons represent emotions to be understood and interpreted the same way by European children and their teachers, as defined during participatory-design sessions. The correlation between the pictures and the concepts they were aiming to represent was tested in the iterations of design, with French and English children (from 4 to 13 years old), and international adults working at the University of Bath. The participants were chosen in order to minimize the cultural differences in recognizing and expressing emotions between the French and English children populations, therefore creating a cross-cultural instrument for use in French and English schools.

SORÉMO PROCEDURE

The *Sorémo* method can be used repetitively during the exploration of a learning environment, in order to capture the user's emotional state at various moments of the interaction with the application, as illustrated in Figure II.7.

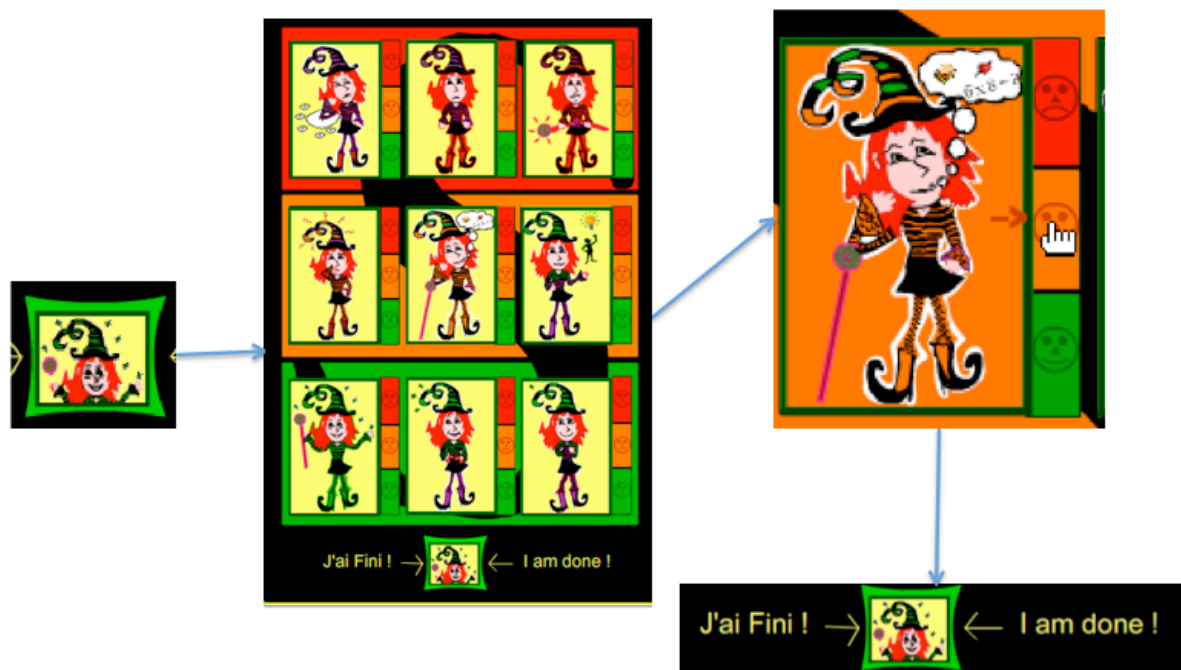


Figure II.7: *Sorémo* procedure

The method is a standalone protocol, that can be used as a single application, or embedded in pedagogical software. Figure II.7 details the procedure when *Sorémo* is part of the *Multipliotest* OLM software: at any moment of software interaction, children can launch a *Sorémo* protocol by clicking on the corresponding button, symbolized by the witch. A *Sorémo* screen then appears and the child can scale each of the nine emotions represented in order to compose an emotional state. For each emotion to scale they can choose the intensity at which they feel the emotion, following a procedure described hereafter. Once they are satisfied that the ratings show “how they are currently feeling”, they can indicate the completion of the task by clicking on the witch button at the bottom of the screen. They are then brought back into the OLM software, as if nothing happened and can continue to use the pedagogical application.

Scaling the 'strength' of an emotion

When asking children to verbally express how they felt whilst playing an educational game, they used expressions to represent the depth or strength of an emotion such as 'I think I feel like that...', 'perhaps like this...', 'a bit like this, but even more like that...'. Therefore a decision was taken to investigate the possibility of capturing emotion 'strength' within the new method. In considering the design and the representation of this scale, it was decided against the use of numbers since they add to the user's cognitive load/effort by having to remember to what the numbers correspond. When asking children in a participatory-design session how they would like it to be represented, they chose to combine the Traffic Light and the Smiley Faces metaphors (cf. Appendix C, section C.2), finding it an easy way to express the strength of their feeling for a particular emotion.

A design choice was made: each of the emotions represented in the instrument would be accompanied by a three-point scale, following the Traffic Light and the Smiley Faces metaphors, and representing the strength of the emotion felt by the user. Each visual picture of an emotion has a Traffic Light scale on the side with visual feedback of the scoring provided by the background colour of the picture, illustrated in Figure II.8.

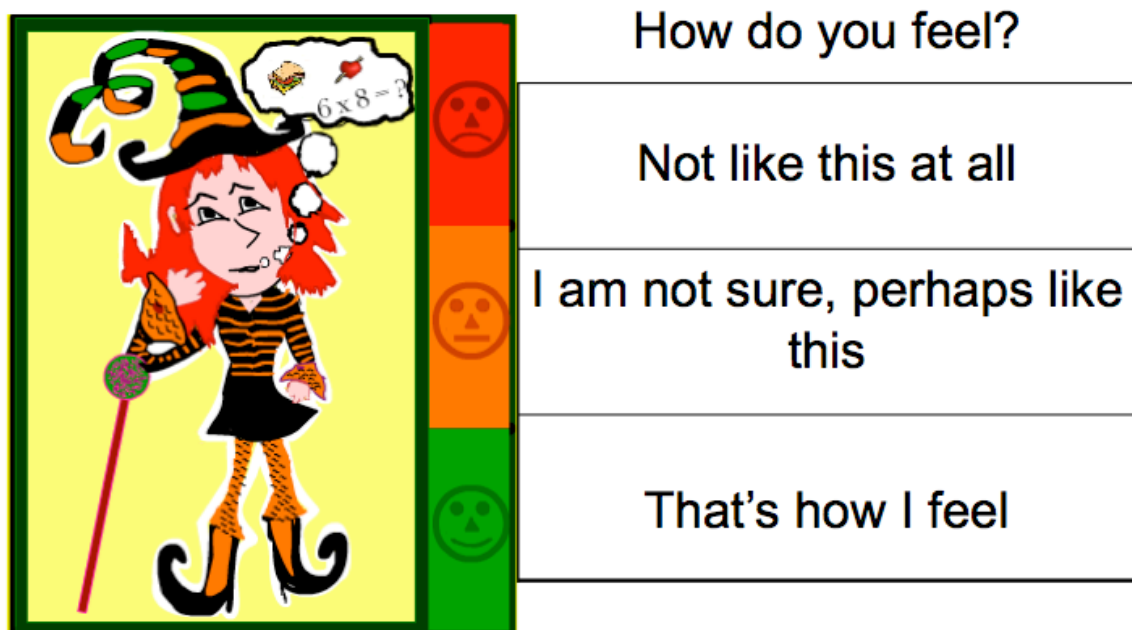


Figure II.8 The Scaling system.

For each emotion, users must ask themselves the following question: “Does this picture represent how I currently feel?”. They then have to select the ‘strength’ score of the scale on the side:

- red and 😞 for “This is not how I feel”,
- orange and 😐 for “I feel a bit like that, but I am not sure”
- green and 😊 for “I feel this emotion strongly”

Although the children participated in the design of the ‘strength’ scale and chose the Traffic Light system, there needs to be a thorough investigation of whether there is confusion, between using the Traffic Lights for both the choice of emotion and its strength, thus adding to the cognitive effort expended. Early evaluations of the new method in use are reported in the next section

Data storage:

Two modes of data collection are available with this instrument: data logs stored in a database using a PHP script; or the creation of an XML data log file for each use of the tool. The instrument records the entire scaling process: the order in which the emotions are scaled, and the succession of scaling that lead to the final emotional state representation.

TESTING THE METHOD’S USABILITY AND USAGE:

The non-verbal self-report tool developed during the PhD to capture children’s emotional states while learning on OLM software was tested in two studies, investigating the usability and potential use of the *Sorémo* method for two different purpose: On the one hand, a study investigates children’s mapping of the emotional value of educational products (cf. Appendix G). On the other hand, a pilot study analysis is presented of how children can use *Sorémo* to represent how they feel while playing, when confronted with the instrument for the first time (cf. Appendix H).

The *Sorémo* instrument, designed with and for French and English children aged seven to eleven years, appears to be understood, and usable for both usability studies, and as self-report instrument to capture a child’s current emotional state. After an exploratory and practice time comprising two trial uses of the instrument, children seem at ease with the emotional representations and the scaling procedure, being more decisive in their answers by using the middle scale only to show a variance in strength of the emotions felt rather than a lack of decision as to the presence of this emotion. There seem to be a developmental difference in how children view affective states: From the behaviours the participants showed to scale their feelings in the studies of Appendixes C and H. It can be seen that some children aged seven to nine still view any emotional state as a unique emotion, which corresponds to Izard’s observations on young children’s representation of emotional concepts

for five to eight years olds (Izard, 1982). On the contrary, all participants aged ten to eleven involved in this experiment only viewed the emotional states as a complex set of basic emotions.

The building of the *Sorémo* method helped gain some understanding of children's capacity to recognize emotions. The results were incorporated in a participatory-design process in order to choose the set of affective responses to children-software interaction, presented in the next section.

2.4.2 DESIGN OF THE A.M.E.R. MODEL

Both the *DividingQuest* and *Multiplotest* software include an affective component, which is composed of activation rules to propose a specific affective response to children's actions. It is used in two contexts:

- As an addition to the pedagogical feedback given during and after an activity.
- As an addition to the visualization of acquired skills in the learner model representation.

The definition of this component's characteristics and artificial-intelligent rules is issued from a work on the application of current theories on learning and emotions for children, and the impact of affective feedback on learning.

DESIGNING THE MODEL: PARTICIPATORY-DESIGN WITH TEACHERS

Designers in the CCI field advocate that involving children in different phases of the design process is vital to ensure that most of the requirements needed are included (Mazzone & Read, 2005). The results from an early study presented in the Appendix E showed that children and adults do not recognize, or understand, affective states at the same level. Therefore, when designing affective educational products, using children from the product's targeted age group should be beneficial to assure a conformity in the desired and actual product use. However, when considering using affect to represent educational feedback, children alone are not enough to design a product of good pedagogical value. Expert teachers are very adept at recognizing and addressing the emotional state of learners and, based upon their observation they take some action that positively impacts learning (Kort et al 2001). The inclusion of teachers working with children of the appropriate age group should help in the definition of an emotional set corresponding to the pedagogical goals pursued.

The participatory-design sessions took place during the development of the two OLM systems, with one teacher from an English school in the UK (when developing *DividingQuest* for English children aged 9-11), and two teachers from a French school (where *Multiplotest* was created for French 7-11 yrs). A first model was undertaken (Girard, 2006) with the English teacher. In order to build a model satisfying the pedagogical needs of both French and English children aged 7 to 11, the

model evolved through iterative design loops, each loop producing a model to be completed and validated by both design partner's parties: the English teacher, and then the two French teachers working together.

The definition of the learning situations to be considered for an EPA's emotional response, in relation to this type of application (OLM drill-and-practice) was performed in the 1st sessions with teachers and children, when designing the pedagogical role, and look and feel of the *DividingQuest* and *Multipliotest*. All emotional responses were created from results in emotional feedback for learning, and observation notes taken from classroom observation featuring the affective interaction between teachers and children. The results of these sessions produced a set of interface scenarios, in the form of sketches of child-interaction with the system.

Constructing scenarios in the design process has for long been recognized as a powerful vehicle in designing interaction. According to Carroll (2000), scenarios have the advantage of being both concrete and flexible, making it easier to cope with the fluidity of the design situation. The scenario represents a particular interpretation of a design situation, but being deliberately incomplete it is also open for negotiation and change.

During the early participatory-design sessions, the teacher design partners used these sketches of student's interaction with the system, and the creation of additional scenarios of interaction. They helped in the definition of responses from the EPA (in French and English). The responses were defined in the form of a written representation of the emotional response in a text box. The teachers were asked to rate the usefulness of each emotional response for the pedagogical role within the software, and provide one (or several) alternative(s) response(s) if needed, with a description of how such responses could be better associated with the pedagogical goal. In the following session, teachers were presented with alternate feedback illustrating the different theories on learning, for each scenario defined. Teachers were asked to make comments on the usefulness of such feedback. The scenarios were portrayed using Microsoft PowerPointTM presentation tools, or low-level prototyping made in HTML/PHP/Flash.

In the next iteration of design, all samples of answers from the first session were compiled in order to present teachers with the alternate possible emotional answers for each situation. The teachers compared the various rationales for using each answer, towards finally agreeing on the definition of a single emotional model.

At first there was some cultural individual differences encountered in relation to child coping strategies when encountering a learning activity. One English teacher was more pronounced in creating a model with a more challenging experience for children with a borderline knowledge, whereas the other two French teachers were adamant that some higher level of encouragement should

be portrayed, to avoid a larger number of children “giving up”. After several sessions on the subject between the design partners and a closer look on the studies showing the results of one strategy or another on children’s behaviour, the model was finally defined, with the following objective: reassuring the low-level and medium-level children in case of failure, congratulating them on their success (with the presence of empathy and satisfaction), and pushing the high-level students to try their best (with emotions such as disappointment when the results were too low).

The choice of dividing the emotional response of EPA’s according to the level of encouragement the child should encounter, was the design partner’s choice, supported by the results of research studies in the literature (D’Mello et al, 2008). The selection of emotions was a long iterative process, where experience and theory were contrasted for the learning benefit of children, in order to create a model of learning at different ability levels. All emotions are supposedly recognized by children of this age, from the results of the literature, and should hold a specific impact on the child’s interaction with the software and learning process, which should help keep the child engaged in the activity and choose the learning strategy of activity that is best for him/her.

In the course of the participatory-design, three design recommendations concerning the definition of affective model of learning with teachers were drawn:

- *Ensure that the understanding of the emotional concepts are the same for all design partners.*

There is a need, when working with an international pool of design partners, to ensure that all partners have the same understanding of what each emotion means. When working with affective computing, the major challenge is to reach a consensus in the definition of the emotions concerned. This matter is amplified when involving people of different cultural and social background, and especially different usual speaking language. The equivalences in intensity and expression of emotions when emotion’s verbal expressions are translated in different languages (here French and English) should not be assumed and should be taken care of carefully. In some cases, an emotion in one language cannot be translated directly into another word, but more of a notion.

- *Provide teachers with applied, descriptive examples of representation of the components to be designed, how they are linked together, and how they will be used in context in software.*

When presenting the different pedagogical scenarios to be used in software, and how the feedback would be portrayed, teachers preferred a visual representation of the situation, in a sketch for example, to a visual demonstration of the software prototype. The situations were described with a high level of details, and needed to be closely linked to the others, with several maps of software flow drawn. Due to the difference in technological expertise, language, and personal choices, it appears that written material was the one most looked forward to by the participants, with numerous illustrative comments.

- *Provide high-level prototypes to teachers to represent software improvements, but also low-level prototypes material for them to sketch out ideas.*

Teachers were keen on using a scenario-based approach to design for the affective responses; form of paper-based sketches of child-software interaction. The use of low-fidelity prototyping helped the teachers to stay on a more general level of design, giving insights about drill-and-practice applications in general, and not only for the particular application domain of the applications to be designed. While we tested some low-fidelity design software during participatory sessions to facilitate the transfer of the results between the two countries, manual paper techniques were preferred. Teachers preferred working on large sheets of paper, with big pens of different colours, and the addition of stickers and notes on sketches situated at the centre of the working space.

EMOTION SELECTION RATIONALE: BETWEEN THEORY AND TEACHER'S PRACTICE

The participatory design approach with teachers as design partners, described in the next section, led to a compromise in emotional responses for the model between theory and practice. For example, a succession of wrong answers was identified by teachers as being a source of moderate to high frustration levels in children of low mathematical abilities. Spurred by Klein et al's advances in reducing frustration (2002), it led to the definition of motivational and empathic responses from the tutor, represented in Figure II.8 as the emotional states comforting and empathic.

Most of the emotions and emotional states selected to be part of the model arise from Kort et al's (2001) set of possible emotions for learning. However, some emotions (sad, amused, inspired) were added from the original affect circumplex (Barrett & Russel, 1998) and in light of the results presented in the Appendix C on the potential capacity of recognition and integration of such emotions by child users. The emotions were structured using the OCC model (Ortony et al, 1998), and presented

to the teachers as an evolutionary decision tree, with children's interaction acting as precursor for specific emotional feedback.

Carver and Scheier (2000) conceptualize affect as input to specific feedback systems. In their view, affect predicts that individuals who feel competent that they can do a task experience positive affect while doing that task, but tend to overestimate their performance and reduce their effort. By contrast, individuals who perceive the task as exceeding their capacity experience negative affect while doing the task; they are likely to underestimate their performance and report increasing effort. Models of affect that view affect as a resource (Aspinwell & Taylor, 1997; Fredrickson, 2001) predict that individuals who experience negative affect during goal pursuit will interpret that feeling as a threat signal and redirect their attention and effort to explore the nature of the threat to deal with it (volition and coping strategies). In other words, Carver and Scheier's (2000) theory predicts that positive emotions reduce effort, whereas resource theory predicts the opposite tendency. Likewise, C&S's theory predicts that negative emotions increase effort, whereas resource theory predicts that they decrease effort.

The state of "awe" is an emotion of high level of arousal and pleasantness. Similarly to admiration, it sometimes leads to inspiration, which in turn can be a motivation to emulate the admired person or otherwise to express what has been learned or discovered (Thrash & Elliot, 2004). The sudden annulment of seriousness and disengagement from prior problem-solving like activity has also been shown to lead to the feeling of "amusement" (Ruch 1993). Happiness is portrayed as a vessel to signal progress has been made towards the realization of a goal (Carver & Scheier, 1990), and that the environment is benign, and safe for both relaxation and exploration (Schwarz and Clore, 1983). The display of such a positive state could therefore lead children to experience more positive emotions towards software use, and help them conceptualize how good their progresses are.

Interest has been identified as the primary emotion that motivates exploration and problem solving (Izard, 1991; Renninger et al, 1992). At first, it may be hard to think of "interest" as a strongly motivating emotion. It is easier to see how anger motivates resistance, or fear motivates escape. However, feelings of curiosity make us ready to tackle a problem or pursue a question until we reach a satisfactory conclusion (Izard, 1991). Interest enhances memory, comprehension, and selective attention (Renninger, 2000; Renninger et al, 1992).

A last emotional state added to Kort et al's (2001) set is "empathy", which has been shown to be potentially useful for keeping the learner motivated (D'Mello et al, 2008; Lester et al, 2000), lowering student's stress level (Prendinger and Ishizuka, 2005), and improving the student's desire to keep working (Bickmore & Picard, 2004). The use of empathy corresponds to Batson et al's view on attribution theory (1995) when confronted with user's boredom or frustration. The EPA's display of empathy portrays an awareness of 'blocked' goals and a willingness to help from the agent. According

to Batson et al, the empathic display may cause the student to understand the agent is attempting to help and will make the student more likely to follow its guidance.

The emotional state of empathy is also used for medium and low level learners when they fail repetitively to answer questions correctly. Following the cognitive disequilibrium theory (Mandler, 1984), at the beginning of the activity, the agent will be more comforting and encouraging so that the student can return to a state of equilibrium, causing a reduction in confusion. However, if the learner persists in failing the activity while trying to answer (not just clicking on whichever button is the closest, or at random), the state of the agent will become more empathic toward the learner's attempts. The agent is acknowledging the child's attempts to reach their goals and direct them out of the negative state of emotions before they give up entirely.

The decision tree of rules defining the emotional responses for each learning situation described in Figure II.8 are articulated according to the following principle: the emotional response of EPAs should be context sensitive and adaptive to the learner's level and abilities in answering the problem to maximize learning. The applications considered are OLM software, where the teachers can define abilities and pedagogical expectations to promote an individualized learning experience (cf. section 2.4.1). Whilst the learning goals can be further individualized by teachers at a later stage, each child initially belongs to a 'level' for the learning of this particular mathematical domain, initialized by teachers before software deployment in classes. This level is inspectable and editable by teachers as the student progresses: they have access to a representation of each student's LM and can modify its content according to the child's progress in class between experiments. Three levels are defined, differentiating students of low, medium, and high-level abilities, as defined in section 2.4.1.

The teachers wished for the EPAs to try and capture the child's attention when presented with instructions as to how the activity works, and mathematical method of answering the problem. For this reason, the EPA's response ranged from polite interest to captivated by the instructions given. Inspiration is also used as the method is revealed in order to mirror the understanding process of the mathematical skill to acquire, and lead children to take some interest in what made the character "get it". Alongside EPA's responses to student's answers of the activity's questions, another state of the interaction was defined: a period of inaction from the student when asked to give an answer. The defined EPA's response was inspired by the cognitive disequilibrium theory (Mandler, 1984), and supported by high levels of recognition of the state of boredom by children corresponding to our final user's age group, as described in the Appendix C.

The principal cause of inactivity in such applications as witnessed by teachers, seems to be boredom. To bring the student out of this negative state of learning, a design decision was made to confront the student with the emotion itself, in the hope that the EPA looking bored would produce a positive reaction from the child. This decision is inspired from the cognitive disequilibrium theory

(Mandler, 1984) and was taken by the team of teachers as a participator-design decision in line with their teaching practices. A second emotional response was associated with this state of inaction: when a longer amount of time has passed without any interaction from the student on the software, s/he is considered as having problems with this particular question. As a system's response to the child-software interaction, the EPA acts as puzzled and prompts him/her to try answering once more, or pass the question if s/he really cannot find the solution.

There initially was another emotional response, corresponding to children passing too many questions, or 'gaming' the system (Klein et al, 2002): the EPA would progressively become angry, and prompt the user to go back to the task at hand. However, the teachers did not reach a consensus on the subject, with some objecting to the 'too negative' aspect of the emotion, which may scare the children away. The emotional state was taken away from the model during the iterative process for this reason.

AFFECTIVE MODEL OF EPA'S SOFTWARE INTERACTION RESPONSES (A.M.E.R.)

Model representation

In Figure II.8 the emotional responses of EPAs are represented according to their valence with a Traffic-Lights System (TL) comprising coloured squares (negative: red, - ; neutral: orange; positive: green, +). The squares at the top of the emotional boxes represent the range in valence of a response. As dynamic representation of emotions in EPAs, the valence of the animations is indeed not constant, but changes with time, and is better represented as a range rather than a single valence position on the axis. For example, when the user is given an instruction, the EPA seems interested, which is considered as a positive emotion of low intensity, represented by a valence going from neutral to positive.

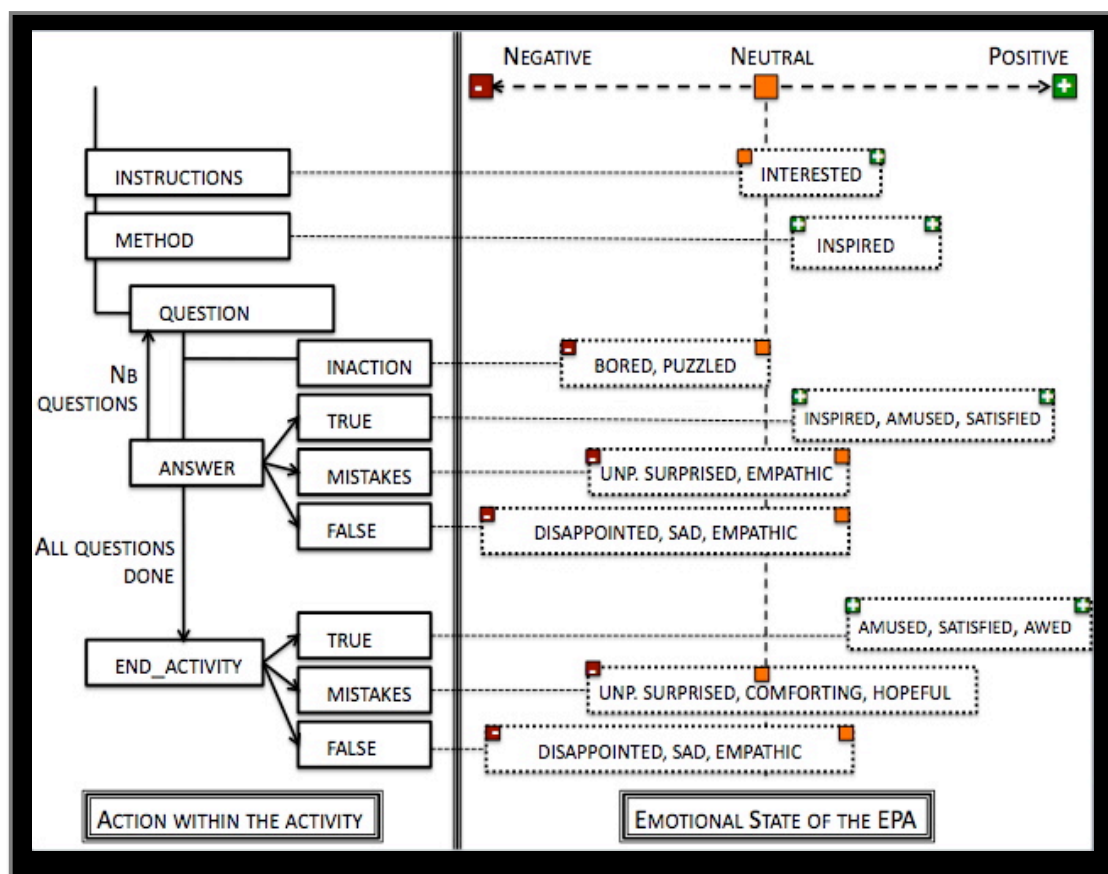


Figure II.9: EPA's affective response model within the OLM systems

A single affective response model was formed for the EPAs of all interfaces used in the evaluations of the research project, due to the similarities in nature and architecture of the two OLM software applications constructed in this PhD, *DividingQuest* and *Multipliotest* : both are mathematical drill-and-practice applications, separated by the number of activities of increasing levels of difficulty, as defined in section 2.4.1.

Each activity is composed of a set of questions the children had to answer correctly in order to “win” the activity. Each question was given a feedback among three possibilities:

- RIGHT: the answer is correct
- MISTAKE: the child knows the process leading to the answer but made an error (numerical error within a long multiplication, or forgot to carry a unit in an addition).
- WRONG: the answer given does not correspond to the questions asked.

At the end of the activity, the right answers are summed up and the resulting number is compared to their individualized pedagogical target, in the LM, which defines the state of the activity:

- WON: they answered a number of questions high enough for the activity to be considered mastered.
- GOOD TRY: they answered enough questions to show some understanding of the underlying concept but their knowledge is not sufficient for the skill to be considered mastered.
- LOST: there were too many incorrectly answered questions, they do not know this concept yet.

EPA's affective Feedback

The EPA responds emotionally to the child's interaction with the activity in the following way:

At first, the EPA is **interested** when presented with the instructions, and points out several important points to pay attention to, and **inspired** along the understanding process of the mathematical skill to acquire, presented in the method of resolution of the problem.

For each question, the EPA progressively becomes **bored** after a period of inactivity on the interface. After an amount of time defined by the teachers for each piece of software, the EPA acted **puzzled** and prompted the child to answer or pass the question it s/he could not answer.

Results of each question brought the following initial responses:

- For a child considered able to do the activity with ease (high level): **inspired** for good answer, **unpleasant surprise** for medium, and **disappointed** for a wrong answer.
- For a child considered borderline to do the activity (medium level): **amused-satisfied** for good answer, **unpleasant surprise** when mistakes were made, and **empathic** when the wrong answer was given.
- For a child considered unable to do the activity yet (low level): **amused-satisfied** for a right answer, **empathic** when the answer is not quite right or definitely wrong.

Those responses change in intensity along with the number of good and wrong answers, reaching the strongest responses represented in Figure III.1, but always keeping in mind the pedagogical strategy of encouragement described in the last section.

Finally, results after completing each question of the activity are a choice between three states "WIN", "GOOD TRY", and "LOSE", according to the number of questions correctly answered and the personalized number they had to get right, taken from their learner-model:

- For a child considered able to do the activity with ease (high level): **amused-satisfied** for a win, **unpleasantly surprised** for a good try, and **disappointed** if s/he failed the test completely.
- For a child considered borderline to do the activity (medium level): **satisfied** when succeeded, **comforting** when a real effort has been made, and **sad** for a failed test.
- For a child considered unable to do the activity yet (low level): **awed** for success, **hopeful** for a good try, and **empathic** in case of failure.

2.5 OLM COMPONENTS DESCRIPTION

The *DividingQuest* was initially developed for an MSc research project to enable investigations as to the impacts of affective EPAs on learning. *Multipliotest* was initially designed as an open-learner model with low interaction between the software and its users, as described in my MSc dissertation (Girard, 2006). The new version, updated during the PhD, enables an interaction with the student's learner model from the children themselves, as well as the teachers.

The pedagogical and expert system components were used identically in the various experiments performed. However, the interaction and access to the LM is the main factor of the experimental designs, and differs in this PhD's work within the experiment itself from one child-group to another. Therefore one can distinguish between three applications used, each application including different user-interfaces to suit the experimental needs:

- In *Multipliotest* 1.0, used in chapter 5, the first version redesigned from (Girard, 2006) to open the learner model to users, the affective feedback is not graphically represented. Software contains two user-interfaces, later referred to as CLM and ILM. The goal is to investigate the benefits of opening the learner model, as opposed to using a closed learner model software:
 - CLM: A user-interface where the content of the learner model is hidden to the child (also called Close Learner Model software)
 - ILM: A user-interface where the content of the learner model is represented using common visualization techniques. The model is available for inspection only.
- In the *DividingQuest*, used in chapter 6, the MSc version of the software (Girard, 2006) was redesigned to suit the A.M.E.R. model (Affective Model of Embodied pedagogical agent's Affective Responses to child-interaction) mentioned in section 2.4.2. Three user-interfaces were formed in order to investigate the benefits of using an affective and coloured EPA to serve as an interaction medium with the learner model content, later referred to as ILM_T, ILM_{PA} and AILM:
 - ILM_T: A user-interface similar to the second interface of *Multipliotest* 1.0 in terms of visualization/access to the model, where the help system and LM view is represented by traditional LM representation tools such as 'skill-meters' and 'boxes'.

- ILM_{PA}: A user-interface; also similar to the second interface of *Multipliotest 1.0* in terms of visualization/access to the model, where the help system and LM view is made through the use of an EPA, which harbours a constantly happy facial expression.
- AILM: A third user-interface; identical to the second interface with the exception of the EPA changing its emotional state according to the affective feedback defined in the affective component.

Previous studies exploring the child's understanding of the Traffic-Light metaphor of learning when applied to pedagogical targets showed the possibility to use said metaphors in ITS for French and English children this age-group (Bull & al, 2005; Appendix B and C). In light of these results, the metaphor was embedded in the design of the application in the form of pedagogical targets, and the visual attributes of the affective EPA (as detailed in chapter 3).

- Finally, another version of *Multipliotest* was built, *Multipliotest 2.0*, and represents three different versions of the OLM, used in chapter 7's experiment, later referred to as u-ELM, AILM, and NLM. It investigates the level of openness of the learner model most appropriate for children to use for better learning and motivation.
 - u-ELM: A user-interface where only children's self-beliefs are inspectable-modifiable by the users in terms of LM.
 - AILM: A user-interface where only the system's view of the learner model is inspectable by children.
 - NLM: A user-interface (also called Negotiated-Learner Model system), where children can compare visually the two representations of the learner model (system's view and children's beliefs), and negotiate the learner model content step by step.

2.5.1 PEDAGOGICAL COMPONENT

As drill-and-practice applications, both pieces of software contain a number of skills to acquire, organized visually in the game according to their level of difficulty and pre-requisite skills needed to attempt each notion.

DividingQuest includes a total of 15 activities representing skills to acquire, grouped by items of 3 into 5 areas of expertise (short/long divisions, problem solving, divisibility, scientific notation).



Figure II.10: *DividingQuest* activity map

In each area, the activities are ranked by level of difficulty (level 1, 2, or 3) and any activity of level 2 can only be attempted once the skill behind the activity of level 1 is well enough acquired by the child, and so on.



Figure II.11: *DividingQuest* levels of difficulty for one activity

Multipliote contains four levels in order to learn how to multiply and to revise multiplication tables:

- Level 1: learning one table of multiplication (numbers from 1 to 10)
- Level 2: practicing short multiplications (1 to 10)
- Level 3: practicing long divisions with one-digit numerators.
- Level 4: practicing long divisions (numbers with two-digit numerators).

Levels 3 and 4 are only accessible to the user once a sufficient ‘level’ (defined by the teacher participatory-design partners and later possible to individualize to each student by their teachers) is reached on the activities of levels 1 and 2.

Figure II.12 illustrates the activity of level 1 and 2: the children have to find the number in the pink box, using the decomposition in units, tens, and thousands to form the resulting number. The question is either finding the result of the multiplication, as portrayed here on the left part, or the second factor of the multiplication, as portrayed on the right part.

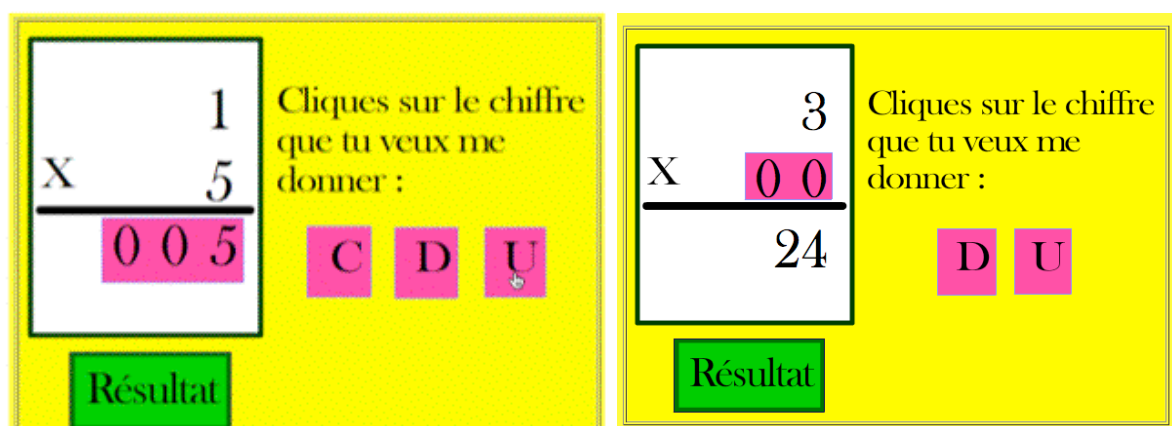


Figure II.12: *DividingQuest* levels of difficulty for one activity

The two OLM software applications constructed in this PhD, *DividingQuest* and *Multipliote*, are similar in nature and architecture: they both are mathematical drill-and-practice applications, separated by the number of activities of increasing levels of difficulty.

Each activity has the same sequence of events: Once they choose an activity, children are presented with instructions and have access to a screen reminding them of the mathematical content needed to complete this activity, allowing practice prior to beginning the activity itself. Once they have begun however, they lose the access to basic mathematical help (e.g. revising the mathematical tables,...), and have to answer a predefined number of questions. For each question, they are presented

with instant feedback on their result. At the end of the activity, the EPA gives them their total result related to this skill and advises them on the next activity to choose.

Three levels are defined, differentiating students of low, medium, and high-level abilities. This approach of dividing pedagogical expectations and emotional support for children was used in class by all teachers involved, and has been exploited in the literature, for example in the AutoTutor project (D'Mello et al, 2008). It can be summarized here in the following way: reassuring the low-level and medium-level children in case of failure, congratulating them on their success, and pushing the high-level students to try their best. High-level students' EPA responses were chosen in order to encourage children to give an answer, rather than playing with the system as a game (Klein et al, 2002). On the contrary, responses for low-level learners contain the highest level of encouragement and supporting behaviour from the EPA.

As previously stated in the literature review, the cost in time of building an ITS with a complete expert system would have been too great for the need of this PhD. Consequently, the pedagogical component includes a competency model that enables the definition of good/wrong answers, but only provides user misconception strategies that would facilitate a low-level discussion of the solution. The learning environments have been built following the constructivist view of learning. They also enable children to work in Vygotsky's (1987) *Zone of Proximal Development* through the learner model visualization and the learning goals of each activity, as described in the learner model component section.

Help screens are available for consultation at the beginning of the activity, and when choosing an activity. A design decision was made not to include help screens during the activity itself, in favour of feedback from the pedagogical agents, as most of the activities are "know by heart" activities. For example, if the table of 2 was available for viewing in the activity that tests the knowledge of this table, the activity itself might become pointless pedagogically. Where the system detects the child having difficulty, it prompts them to use help. The questions offered increase in difficulty as the child progresses successfully, similarly to the *Subtraction Master* application (Bull & McKay, 2004). Where there are problems, the system guides the child through the subtraction process. If a possible misconception is detected in the *Subtraction Master*, or a lack of knowledge of a specific concept, further questions are selected to elicit data on the likelihood of the child holding that misconception. In this PhD's work, software reaction to child's misconception is limited to focussing on a specific table, or offering to view the method linked to the concept. For example, in *Multipliotest*, the questions testing a child's knowledge of all multiplications from 1 to 10 will focus on the tables the child has been found to have trouble with.

2.5.2 LEARNER MODEL COMPONENT

The learner model component of each piece of software is composed of the following parts: user's characteristics and user-interface preferences, a set of pedagogical goals given by their teachers, children's level of acquired skills for each activity, and finally children's self-belief as to their knowledge of each skill. The latter is not typical of OLM software, and will be rationalized and detailed further in this section.

OLM DESCRIPTION: SMILI[©] FRAMEWORK

According to Bull & Kay's SMILI[©] framework (2007), discussed in details in section 1.4.3, the purpose of any OLM application can be classified according to two concepts:

- Context and evaluation: how does the OLM fit into the overall software interaction?
- Centrality: how central is the openness of the LM content to the system's aim: how do the learners make use of the model, and how does it fit into their broader learning activities?

Broad descriptors: Centrality, Context and Purposes

Table II.3 compares the applications used in the experimental chapters (5 through 7) in terms of context and centrality of the models, and design of the experiment:

Table II.3: SMILI[©] OLM Framework used to describe the *DividingQuest* OLM software: the general issues.

	Chapter 5 CLM, ILM	Chapter 6 ILM _T , ILM _{PA} , AILM	Chapter 7 u-ELM, AILM, NLM
Centrality	Openness of the model	Presenting the model: EPA	Types of Interaction with LM
Context	Multiplications	Divisions	Multiplications
Evaluation			
Design	Between subject, single session on software	Between subject, single session on software	Between subject, repeated Measures
Size	Medium	Small	Extensive
Participants	14 per condition, 28 in total age 9-11	9 per condition, 27 in total age 9-11	15 level CE2 and 15 level CM1 per condition, 90 in total age 7-9 for level CE2 age 9-11 for level CM1
Data	Mathematical pre and post test, interaction logs, interview, questionnaire	Mathematical pre and post test, interaction logs, interview, questionnaire	Mathematical pre and post test, interaction logs, <i>Sorémo</i> self-reports, child and teacher interviews questionnaires

The *DividingQuest* and *Multipliotest* applications have a teacher's interface with personalization of individual and class learning by access to children's learner models (for individual coaching, to check the class's progresses on a particular notion). Teachers are given access to the child's model with edition rights, therefore increasing its *accuracy* by allowing not only a systematic approach to model updates through the software, but also some rectifications according to the child's progress between software use. The OLM helping him/her *monitor* each child's progress and *plan* future learning sessions on the subject. However, while teachers effectively used their interface when software was displayed in schools, this PhD focuses specifically on children's reactions to and use of OLM.

The first experimental study, presented in chapter 5, compares children's behaviours and achievement on software with inspection rights to the LM (ILM) or no access to the LM (CLM).

In the experiments of chapter 6, the extent to which children are willing to use the content of the LM for learning and how it changes their interaction with software is investigated. The three interfaces of the *DividingQuest* software (ILM_T, ILM_{PA}, and AILM) are therefore centred around enabling and recording children's inspection of the learner model.

At the beginning and the end of the learning session with software, children use a smiley-o-meter (Read et al, 2002) for each separate knowledge component in the system. These self-assessments form the student's belief measures in the learner model (later referred to as children's self-identifications), and help gain a learner model more *accurate* to the child's level: children using the facial/bodily expression and/or TL colour cues to express their understanding of a particular concept. The results of these assessments can then be compared to the system's LM content at the time of data analysis, and a level of accuracy can be drawn (Kerly & Bull, 2006).

In order to investigate children's use of OLM applications, this study introduced in chapter 7 proposes three types of interaction with the model (inspection of the system's view; modification of the user's view; and negotiation between the two views):

- u-ELM: this OLM was designed with a main focus on the issues of control of the OLM parts. Accuracy and reflection are also at the centre of the evaluation goal, as children are total 'masters' of their view of the OLM. While the expert system and teacher's view of the child OLM are ruling the pedagogical level of difficulty for the activities, children can define their view of the LM with self-identifications, modify this view at the end of each activity, and rely on it to plan their activities and navigate in the interface.

- AILM: this OLM system holds the same general purposes and centrality of OLM as the one described in chapter 6. It is centred on using the inspection of the learner model to facilitate user's reflection, help them plan their activities. It also aims at helping them navigate better through

the software in terms of choosing activities content and purpose. The accuracy of the model is also explored using the self-identifications of students' belief as pre and post tests, as well as when they initiated it while using the software.

- NLM: The primary goal of this interface is to improve accuracy of the learner model by using a negotiated model, therefore producing the best opportunity for learning. It also aims to help children plan their learning session according to their currently mastered skills as well as the ones they wish to acquire.

Each interface includes the same mathematical content, EPA representations, and emotive responses to interface (u-ELM, AILM, NLM) interactions similar to the description of *Multiplotest* 1.5 used in the study presented in Appendix J. However, a new EPA 'Moije' has been included to represent children's beliefs, of their learner model and used in the interfaces along with 'Multiplio', the learning companion EPA representing the expert/teacher view of the learner model, as illustrated in chapter 4. This helps keep two sets of belief measures, which is important for representation as both are taken into account during an interaction. The EPA created, named 'Moije' (literally 'Me I', phonetically similar to the French word 'Mage'='magus') by the participatory-design children, is a stylized, simplified, dynamic representation of a smiley-like basic EPA model, as developed further in the section "System's and Child's view of the LM".

The centrality of the openness of the learner model to the interaction will be evaluated from the user's actions: some children will have a minimum level of use (when the system shows an overview of the model to choose an activity), to only using it at the system's initiative (when the student is having problems, and immediate feedback of answers), or extensively (watching the model update after an answer or an activity). In chapter 7, the difference is also linked to the content of the experimental condition: In u-LM, the users are forced to use it to a greater extent as they are prompted for self-identifications within the activities. The same can be said for the NLM interface, with the negotiation process. In s-LM however, the (non) use of the user model remains the initiative of the user.

Some results from the literature show that while allowing peers to access (parts of) the class's or other students' learner model can be beneficial for some (Kerly & Bull, 2006), it can also be detrimental to the learning of others (Bull et al, 2005). For this reason, the *collaboration* and *competition* aspects of OLM were not investigated here, and left for future work. We simply cannot tell how the results described in this thesis could be transferred to a collaborative application or one with an OLM (partly) shared by peers.

Elements of the OLM application:

The framework used to enable comparisons of OLM systems, and represented in table II.4, is inspired from the SMIL[©] framework. A cross in a cell shows that the corresponding item is present in this software interface.

In this section, the content of, and access to the LM is compared across all the interfaces bearing an open learner model, grouped as follow:

- ILM: an inspectable learner model, that does not show differences in affective states in the system's view of the LM during inspection. This corresponds to the interfaces ILM in chapter 5, and ILM_T, ILM_{PA} in chapter 6.
- AILM: inspectable learner model, where user feedback and LM content (system's view of the LM) are represented following the AMER model, introduced in chapter 3. It corresponds to the AILM conditions in chapter 6 and 7.
- u-ELM: software where children can only access to their own representation of knowledge acquisition, the system's view of the LM staying hidden. This corresponds to the u-ELM condition in chapter 7.
- NLM: open-learner model where the learner model content is built as children used software, with a negotiation between children's and the system's view of the child's level for each skill, to edit the underlying LM. This corresponds to the NLM condition in chapter 7.

Table II.4: Adaptation of the SMILI© OLM Framework describing in more details the elements of learner model from each pieces of software.

Elements of the OLM	Properties	Description			
		ILM	u-ELM	AILM	NLM
1. Extent of model accessibility	Complete				
	Partial	X	X	X	X
	Knowledge level	X	X	X	X
	Knowledge	X	X	X	X
	Difficulties				
2. Match underlying representation	Misconceptions				
	Learning issues				
	Preferences	X	X	X	X
	Other 'users' LM				
	Affective Model		X	X	X
3. Access to uncertainty	Similar system LM	X		X	X
	Similar user LM		X		X
4. Role of time	Complete				
5. Access to sources of input	Partial				
	Previous				
	Current	X	X	X	X
	Future				X
	Complete				
6. Access to model effect on personalisation	Partial				
	Affective State	X	X	X	X
	System	X		X	X
	Self		X		X
	Others				
7. Presentation	Complete				
8. Access method	Partial				
	Textual				
	Graphical	X	X	X	X
	Overview	X	X	X	X
	Target detail	X	X	X	X
9. Flexibility of Access	All details				
	Support to use				X
	Affective Feedback		X	X	X
	Inspectable	X	X	X	X
	Editable		X	X	X
10. Access: initiative comes from	Addition				
	Student persuade				X
	Negotiated				X
	Complete				
	Partial	X	X	X	X
11. Control over accessibility (to others)	System		X		X
	User	X	X	X	X
	Peer, instructor				
	Complete				
	System, other...				

The CLM interface does not belong to table II.4 as no item of the model is visible or accessible for inspection or edition, at any time of software use. This is due to the nature of a Closed Learner Model application, where the model is, by definition, hidden from the learners.

ELEMENT 1: *Extent of model accessibility*

All interfaces with an open learner model give partial access to the learner model with inspection of the knowledge level, with some access to the difficulties children encountered during the questions and summary of activity feedback. However, they do not allow children to inspect a list of their misconceptions outside of an activity itself, unlike other OLM systems such as Flexi-OLM (Mabbott & Bull, 2004).

Children also have access to ‘knowledge modelling’, represented by the ‘lecture’ LM view. The structure of this representation remains simple, and is similar to the representation of concepts as presented in the screens of activity selection. Due to the nature of the drill-and-practice software, the notions are clearly identified, and the hierarchy between concepts in terms of levels of difficulty are similarly represented.

For the next sub-section, learners can give their preferences concerning the user interface and elements manipulated in the activities, but not on the content, or difficulty of the activities provided.

A last sub-section has been added to the SMIL[©] framework in this element: the possibility to access the information of the affective model of the applications.

ELEMENT 2: *Match underlying representation*

All systems offer a visualization of the user model similar to the underlying representation, using the EPA as interaction means to simplify the content.

A last specificity of OLM will be discussed here: the use of a second view of the model corresponding to children’s self-identifications when using the *DividingQuest*. In both pieces of software, children were asked as a prelude and end to the learning session to evaluate their current level of knowledge acquisition for each concept manipulated. These ‘self-identifications’ are no longer available after completion. The software’s view is first initially filled by teachers, considering children’s skills on each concept. It is then modified by software use and the corresponding results to activities.

ELEMENT 3: *Access to uncertainty*

As in Bull & MacKay's *Subtraction Master* (2004), children are considered too young to fully comprehend the notion of uncertainty when applied to learner model, and therefore did not have access to this aspect of the OLM. This means that the representation of the learner model does not include percentages or levels of certainty, about the model's accuracy integrated to the child's current level. The representation of the LM therefore aims at accurately representing what children currently know.

ELEMENT 4: *Role of time*

Interfaces ILM and AILM have access to the current view of the learner model, however they do not have access to learning history, nor to the anticipation of future states, these notions considered borderline for the cognitive development of the youngest users by the teachers, ie: too many children would not be able to, or would have too many difficulties in understanding this concept, making this feature more distracting and confusing than useful for the interface interaction. This was considered as a potential bias in the experiments and would have to be investigated separately.

Each software has access to the current view of the OLM (either a view of their self-beliefs, the system's beliefs or both views to the interface). Only the Negotiated Learner Model (NLM) has access to the future states, during the negotiation process only: when an activity is finished and the final results appear a negotiation can begin if the system and child's view of the results level diverge, using smiley-o-meters. In this case, the user has access to the current level and the system shows what the future state will be. After the negotiation, the user has again only access to the new current state.

ELEMENT 5: *Access to sources of input*

There is no complete access to the data input to children: some data is entered by teachers during software initialization, and other data recalculated by the system along with the software along with the software use. However, children all have a partial access with the self-identifications to the self learner-model at least on their own view. All systems have full access to the 'self' part of the OLM input, and no access at all to 'others' (i.e.: the teacher's view). This does not modify the LM in system's view, but adds contextual information, later used to test the model's accuracy.

ELEMENT 6: *Access to model effect on personalisation*

There is no access to the model affect on personalisation. Like in the *Subtraction Master* (Bull & MacKay 2004), it was decided by participatory-design teachers not to include this aspect to avoid distracting children with it.

ELEMENT 7: *Presentation*

DividingQuest and *Multiplotest* were designed to support the user on OLM issues only with the help of the EPA for interaction, without any additional information. The presentation is more focussed on graphical information than textual due to the age of the children and their literacy level. The information presented as a ‘related concepts’ or ‘lectures’ view of the model of each concept to acquire, as well as targeted detail representation per main skills to acquire. No help system of use other than the EPA was offered as support for the interaction considered (and tested in usability) explicit enough as is.

ELEMENT 8: *Access method*

Children in ILM and AILM interfaces have only access to their learner model by ‘inspecting’ it, they cannot edit, add information, ask for a test to persuade the system or negotiate the results. If they want to persuade, they have to do the activity corresponding to the learner model element, and the results corresponding to their belief.

This is the main factor studied in the evaluation of chapter 7: In u-ELM, the students can inspect only their view of the learner model, whereas in NLM and AILM they can also inspect the system’s view of the learner model. All systems can only edit their own view of the learner model through self-identifications, they cannot add evidence to the system’s view nor persuade them to change a model by requesting a test. However, unlike in u-ELM and AILM, NLM offers the possibility to negotiate the content of the learner model within an activity. Similarly to the system in chapter 6, in order to persuade the system of one’s own belief, one has to choose the corresponding activity and ‘prove’ to the system that their view is the right one by successfully answering the questions.

There is another part of the learner model partly accessible to the users: the self-identifications of children’s emotion state, given by the completion of the self-report *Sorémo* method presented in chapter 3. This part of the model is editable for all interfaces when children are using *Multiplotest*, but cannot be inspected later on by the children. However, in the teacher’s view, not under investigation here, teachers have access to this part of the model, with a targeted or general visualization of the evolution of children’s emotional state during a learning session.

ELEMENT 9: *Flexibility of Access*

Software offers some flexibility of access, restricted to the views authorized for access : provides overviews on request, further details on certain issues.

ELEMENT 10: *Access: initiative comes from*

Due to the reflection and planning central goals, the access to the learner model can be initiated not only by the system but also by the user. This is also essential to enable the negotiation process on the NLM condition.

ELEMENT 11: *Control over accessibility (to others)*

Finally, there is no control over the access to other actors of the learning process (peers, parents, etc...), the learner model in every interface was stored via the internet on a database of the university, the *DividingQuest* and *Multipliotest* being designed for online use in classroom settings.

USER'S CHARACTERISTICS AND UI PREFERENCES

In this part of the learner model, all information about the child's personal characteristics and software personalisation are recorded: age, gender, colour/learning disability, experimental condition tested, and then user's UI visualization and item preferences. For example, in the *DividingQuest*, children can choose a set of pictures representing their every-day life interests in order to use familiar/favoured notions in the problem solving activity.

PEDAGOGICAL GOALS AND EXPECTATIONS

When registering a child into the system, his/her teacher attributes him/her a general level corresponding to the child's current knowledge, and future skill acquisition expectation level. The last information entered corresponds to the type of feedback deemed more suitable to work with this child (reassuring, praising, challenging/pushing or the best). Such information defines a first set of children's achievements level for each activity, and type of pedagogical/affective feedback to be given by linking the information from the various components. First registered when adding the child to the system, the information is still editable during the period of software use by teachers. This late edition will not be investigated in this PhD, the research focussing on children's use of OLM applications, and not learning regulation by the teachers.

SYSTEM'S AND CHILD'S VIEW OF THE LEARNER MODEL: LEARNER'S ACHIEVEMENTS

System's view of the LM:

For each activity, the results concerning each question answered by child-users (question + answer + misconception) are registered within the software, and help in the definition of the child's level for this activity, as well as future expectations. For each skill defined in software, a level of child's achievements is associated (good, medium, and bad) according to the values initially set by teachers, as well as the results of software use.

As an example, Charlie is considered of medium-level in his apprenticeship of tables of 2 and 7 with a teacher's expectation of "average" (he doesn't have to answer correctly every single question to be considered as mastering the activity). In the French system for the learning of multiplications, the tables of 1, 2, and 10 have to be mastered perfectly before attempting to learn the other tables. Consequently, while working on knowing better the table of 7, it is absolutely necessary that he knows the table of 2. The expectations set by default in the LM component for a child expected to be "average" to master the table is set at least 8 correct questions out of 10 for the table of 2, and at least 6 correct for the table of 7. In order to push Charlie to learn his table of 2 better, the teacher can choose to modify the number of questions answered correctly at 9 or even 10 correct questions out of 10. In this case, the system's view of the learner model will go from "good" to "medium", as the expectations are not fulfilled. Once he masters this multiplication table, the expectations for the tables considered more difficult will automatically arise.

Children's self-identifications: Child's view of the LM

At the beginning and the end of each software use, children are asked to rate their current skills acquisition using a 3-point Likert scale (good, medium, bad). The history of child's self-beliefs about their own level is registered and used when solicited by the OLM main controller when a comparison must be done with the system's view of the model of chapter 7. In some of the experiments described in this PhD, the child can view and/or modify the content of these values during software use.

2.5.3 USER-INTERFACE COMPONENT: INTERACTING WITH THE MODEL

THE OLM APPLICATIONS STORY FRAMEWORKS:

As previously stated, two OLM systems are used in this PhD work, build for two specific mathematical domains: the apprenticeship of divisions, and multiplications. The user-interface of both application differ significantly in terms of design and perception of activities, as illustrated in Figure II.13:



Figure II.13: *DividingQuest* and *Multipliotest* software interface

The *DividingQuest* application is an adventure and fantasy drill-and-practice game, whose storyline aims to “change the world by succeeding mathematical challenges”. The Embodied Pedagogical Agent embedded in two of the game’s interfaces is a fantasy character (the witch ‘Séléna’) in line with the “*DividingQuest*” theme: inspired from the Lord of the RingTM book and RiskTM game, the users progress one activity at a time, freeing regions taken by the big bad Dragon, showing the townspeople of corresponding areas how to forge different pieces of armour to defend themselves with mathematical knowledge. In *Multipliotest*, the game-like scenario is less apparent, as the interface merely shows a set of skills to acquire, at the end of which the child is proclaimed “Multiplication Master” and given a diploma.

REPRESENTING THE LM IN SOFTWARE:

Both pieces of software include an adaptive help to visualize and use the LM, using either common visualization tools such as “skill-meters” and ‘boxes’, or proposing an EPA as an interaction medium with the LM, following the results on children’s appropriation and preference of LM visualization techniques, presented in chapter 4.

Multipliotest offers two views of the LM: First, a “related concept” view under the form of an activity to choose, and using “boxes” as main visual representation tool, with EPAs representing either the system’s or the child’s view (or both in NLM) of the model, using Color-o-smileys, according to the experimental condition evaluated. Last, a “lectures” view of the different skills manipulated in the software in the form of a list of skill-meters for the children to use in their self-identification selections (as a prelude and an end of the learning session of the application, and whenever they wished to change the value of the icons selected during the learning session).

Skill-Meters representation of the LM content:

In the interfaces of inspectable learner models with “traditional” LM representation tools, the skill –meters were represented using skill-meter bars for the *Multipliotest* software (condition ILM in chapter 5), and pedagogical targets for the *DividingQuest* software (condition ILM_T in chapter 6).

For all the interfaces with an affective EPA, a special kind of Likert scale is introduced: the *Colour-o-Smileys*. Smiley-o-meters, part of the Fun Toolkit (Markopoulos et al, 2008), have been found a reliable means to gather opinions from children (Read, 2008). In the child-computer interaction community, the Smiley-o-meter, a smiley-based survey instrument, is a popular method to evaluate fun with children. Surveys based on face or smiley scales know a long tradition in medical research for measuring pain (Bosenberg et al., 2003; Vorderer & Bryant, 2006) or anxiety and in psychosocial situations for assessing relationships (Denham & Auerbach, 1995). The Smiley-o-meter was introduced into the Child-Computer Interaction community as one instrument of a methodological ‘Fun Toolkit’ for the evaluation of fun in children’s product interactions (Read & McFarlane, 2006; Read et al, 2001). The roots of the Smiley-o-meter are positioned within the family of 5-point Likert scales (Read et al, 2002). Response options (be it numbers or categories) are replaced by smileys if it concerns attitude information related to for instance fun, preferences or user experiences (Read & McFarlane, 2006). The ordinal response options are scaled with higher values reflecting a more positive attitude. Although the Smiley-o-meter proved valid for the measurement of fun when administered with children, its use by young children in order to express their opinion still remains to be investigated.

Contrary to the smiley-o-meter scales usually used for usability studies with children, the design of the design of our scale includes not only facial expressions, but also the Traffic-Light metaphor to represent stages of learning acquisition, as illustrated in Figure IV.14:

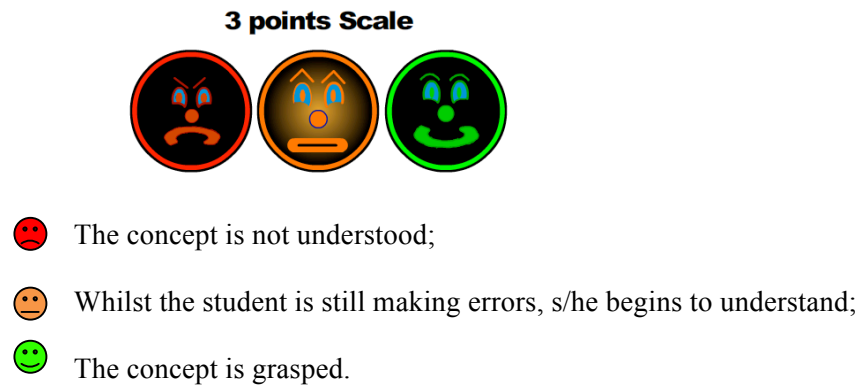


Figure IV.14: The 3-points scale Colour-o-Smileys separates learning in three states.

Boxes

The second type of LM representation tool used in every interface of the *Multipliotest* software in chapter 7 is “boxes”. From the observations and results of section 4.2 and a pilot study on children’s preferences in EPA’a appearance presented in appendix E, two EPA icons have been designed in participatory-design with children and teachers from the same age group, to be used in the representation of the LM in *Multipliotest*.

The character representing the system’s view, “Multiplio”, is represented on the left in Figure II.15, while the symbol chosen for the representation of the child’s beliefs, “Moije”, is represented on the right.

One box represents one skill to acquire: either a multiplication table (here table of one), or a multiplication method. It contains the description of the skill (text/numbers), and according to the experimental condition, the system representation (*Multiplio* symbol), and/or the user representation (*Moije* symbol) of the user’s acquisition of said skill, as illustrated in figure II.15:



Figure IV.15: Example of “boxes” in the “concept map” view for each interface

This example represents a hypothetical learner model representation for the knowledge of the table of one. In the NLM interface representation, the user rated his/her knowledge of the table as “concept not acquired” by selecting the red and unhappy symbol for the “Moije” character. The system evaluation of the child’s knowledge, however, considers that the child knows his/her table of one perfectly, as illustrated by the “Multiplio” symbol in green, portraying a happy expression. The child is always presented with the “Multiplio” character as a representation of the system’s view, on the left of a “Moije” character that represents the child’s view on the model.

Related concept map view

In the *Multipliotest* software, the participatory-design teachers preferred a related concept map representation by encapsulation of units instead of the concepts being linked by arrows or lines, children this age being more used to this type of representation in school activities. All ten multiplications tables were therefore encapsulated into a “level 1: one multiplication” box, while there was a single box for “level 2: all multiplications 1 to 10”, and one box for “level 3: multiplying big numbers” with two sub-boxes, namely “multiply a digit by a number”, and “multiply two big numbers”. The order of the boxes followed the Latin sense of reading (from left to right, and top to bottom), in such a way that all fit into a single screen, without the need to scroll (either horizontally or vertically).

Figure II.16 illustrates the related concept map view for the interfaces NLM from chapter 7:

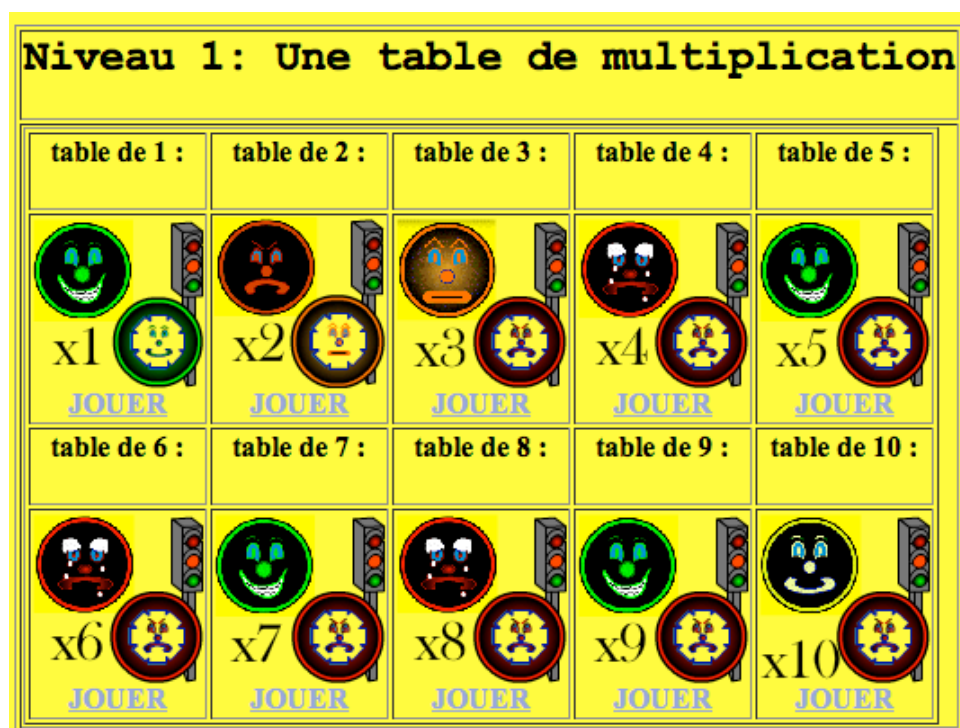


Figure II.16: “related concept map view” for NLM interface, activity of level 1

Lecture view

The “lectures” view lists off the different skills manipulated in *Multipliotest*, separating them by the level of the game they belong to. Each skill is represented by a textual description, followed by a *Colour-o-smiley*. For each skill, the weight of the scale points to the user’s current identifications of his/her knowledge, as registered in the application from previous interactions.

INTERACTING WITH THE LM

This section describes the interaction medium proposed to children to inspect and edit the LM in the *Multipliotest* software. This corresponds to the interfaces AILM, u-ELM, and NLM in chapters 5 and 7. The process is similar to the AILM interface for the AILM, and ILM conditions of chapter 6 in the *DividingQuest*.

Children first perform a series of three mathematical questions for the pre- session mathematical test, as described earlier in Figure II.12. The next step required is the completion of a self-identification questionnaire for the user to give their own beliefs on each mathematical skill encountered. The interface, illustrated in figure II.17, is used in the pre- and post- session self-identification screens. It is also accessible in the u-ELM and NLM conditions to modify their view of the LM at any point during software use.

The screenshot shows a yellow background with the title "Quelles tables de multiplication connais-tu?". Below the title are ten boxes, each representing a multiplication table from 1 to 10. Each box contains a "color-o-smiley" meter with five colored circles (red, orange, yellow, green, blue) and a corresponding number of dots below them. The meters for tables 1 and 9 are fully green, indicating the highest level of knowledge. The meters for tables 2 through 8 are partially green, indicating intermediate levels of knowledge. The meter for table 10 is partially green, indicating a lower level of knowledge. A green button labeled "Envoyer" is located at the bottom right of the interface.

Figure II.17: Part of the self-identifications in the AILM, u-ELM, and NLM interfaces

Children have to choose, for each mathematical skill represented in the learner model component of *Multipliotest* what level they believe they have currently achieved for the apprenticeship of this skill. In Figure II.17, the child considers that she knows her table of 1 and 9 by heart, as she selected the highest grading point in the color-o-smiley meter.

The children, once on the interface can choose to inspect the learner model, and have then access to the interface presented in figure II.15 (NLM condition represented). In the event that children want to modify their representation of the LM during the game (conditions u-ELM, and NLM), they can select the “Edit” button, illustrated in figure II.18, and follow the same process.



Figure II.18: Edition of the self-identifications, u-ELM interface

Once on the game interface, they can choose an activity and then answer each question that compose that activity, as described in section 2.5.1. In the AILM condition, some graphical feedback is given to children about their performance on each question, using the EPA representing the system’s view of the LM: Multiplio.

Figure II.19 illustrates the feedback given for a wrong answer of activity of level 1:

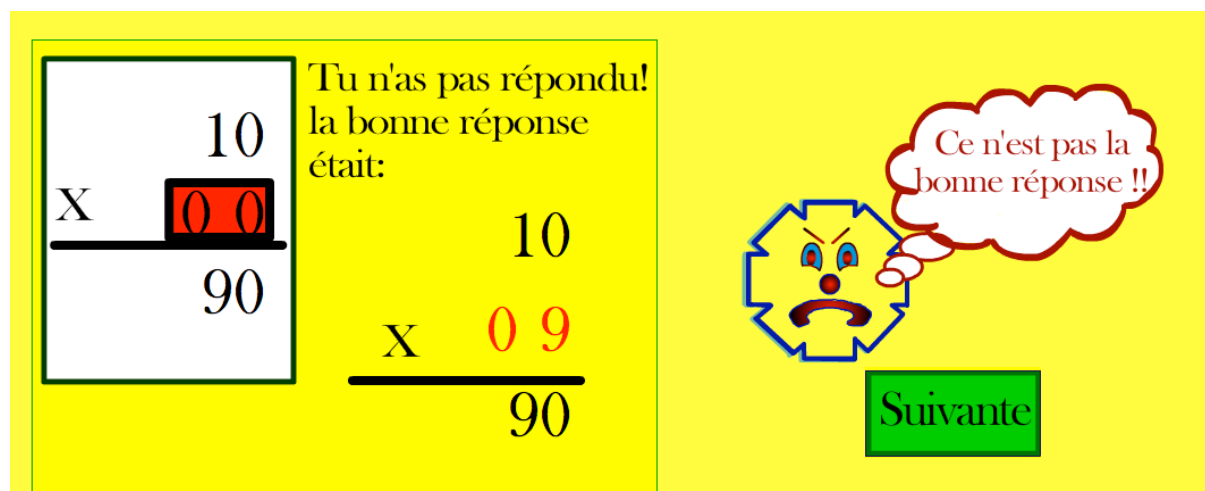


Figure II.19: Multiplio feedback for the AILM interfaces in case of wrong answer

When the activity is over, some graphical feedback is again given to children using Multiplio, as illustrated in Figure II.20 for an activity graded as a “good try”, as detailed in section 2.4.2:



Figure II.20: Multiplio feedback for the AILM interfaces in case of a “good try” activity

In the u-ELM interface, the Multiplio interface does not appear, as illustrated in figure II.21. However, the user can give his or her own judgment about the feedback to be given for this question, using color-o-meters.

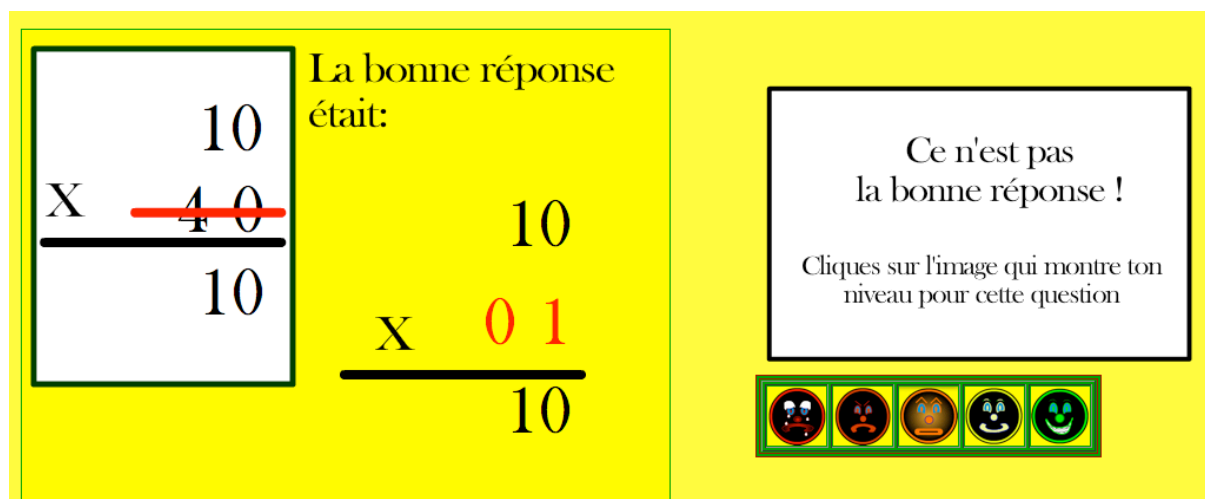


Figure II.21: u-ELM interface, give your own feedback for a question

Finally, in the NLM interface, the user can enter a process of negotiation with the system to define the level of feedback for one question, as detailed in figure II.22. In this case, the user is first asked what he thinks of his performance, as portrayed in figure II.21 and used in condition u-ELM. Another screen then appears (figure II.22), giving the feedback deemed correct by Multiplio.



Figure II.22: Negotiation process, give your opinion for NLM interface

And the user can either leave his own belief as it currently is, or rate the color-o-smiley another time. The two views do not necessarily need to end up identical, and will both be represented in the progress bar during the experiment, as illustrated in figure II.23.



Figure II.22: Progress bar for NLM interface

For the feedback given to the end of an activity, the negotiation process is more extensive. When there is a conflict between the child's and system's views of the learner model, identifiable directly by the graphical representations of the Moije et Multiplio characters, the child can try to "persuade" the system: he or she is then given a sample of questions that were deemed unknown by the user to complete. According to the child's answer, the estimation of the system's view is modified. Once gain, there isn't necessarily the need for total agreement between the two views at the end of the process.

2.5 CONCLUSION

In this chapter, two OLM educational software, concerned with the apprenticeship of mathematical concepts, are introduced. Each software, presents different user-interfaces which, while identical in terms of mathematical content and student's achievement expectations, differ in how the information about the learner model is presented to the child, and how s/he can access/modify such information. The following research questions were addressed to some degree:

RQ1: What theories of emotion can be followed to produce a model of EPA's affective response in OLM software suitable for mathematical drill-and-practice applications?

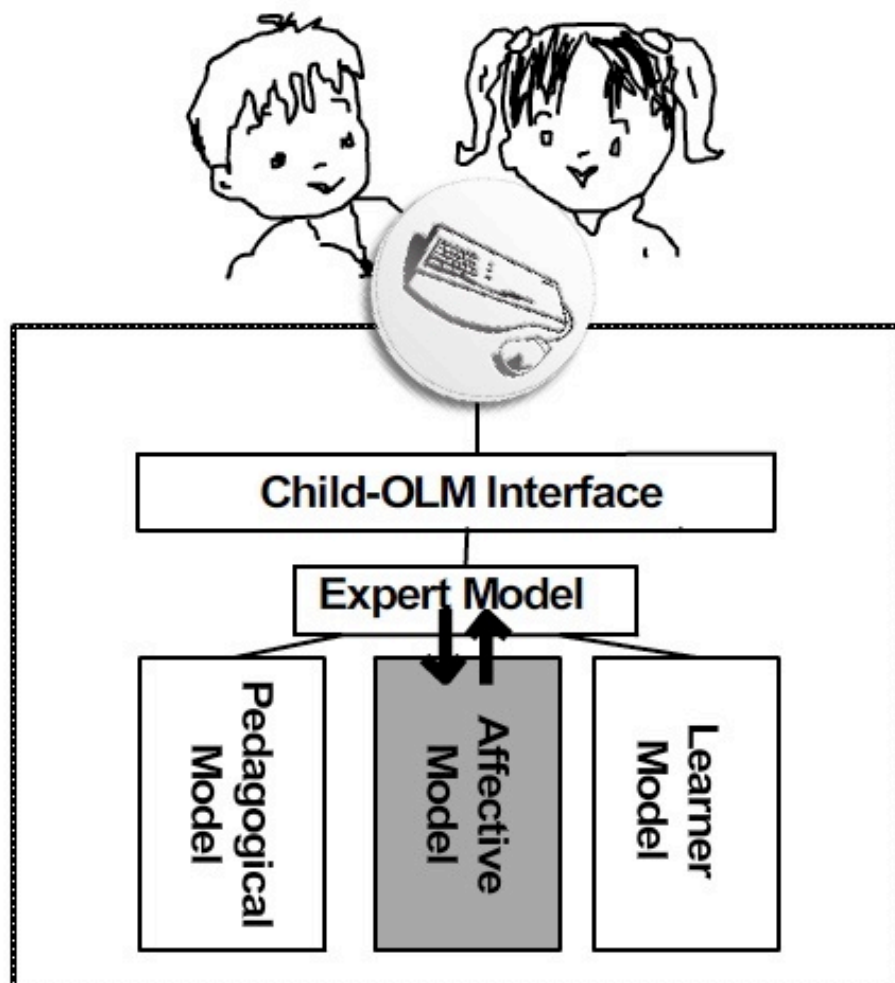
RQ3: What is an appropriate representation of a learner-model (LM) for a child to interact with?

RQ4: How can user-centred and participatory design techniques be used on children and their teachers to represent the Learning model components of OLM tutoring systems for children?

A model of emotions, A.M.E.R., developed with teachers as participatory-design partners, and based on current research in affective computing, was presented. The model has been especially designed for drill-and-practice mathematical OLM systems with child users aged seven to twelve years. The model, once embedded in the *DividingQuest*, will be tested in terms of benefits for learning and motivation when compared to other types of feedback to child-interaction with the environment (see the experiment presented in chapter 5).

Chapter 3 will now describe in more depth the affective component of the OLM in exploring how affect can be added to pedagogical feedback given to child-users in order to increase learning and motivation.

Chapter 3: Emotion, Children, and Learning



3.1 INTRODUCTION

This chapter is concerned with the affective components of OLM educational applications: where can affect be embedded in OLM tutors and for what potential gain? Research studies in educational technology have been attempting to define the relationship between student's affective states and their use of intelligent tutoring systems through different means (Kort et al, 2001; Chaouachi & Frasson, 2010; Irroyo et al, 2010).

The first aspect investigated concerning affect is the detection of a child's emotional state, and its evolution through software use. The second lies in children's reaction to emotional feedback in a pedagogical context, with the following question: does the simulation of emotions through digital characters at strategic moments of user-interaction within educational software has an impact on children's motivation level and learning outcomes? The results show there is evidence that affective EPAs (EPAs representing different states of affect according to user's achievements) can help increase the level of motivation for some children in traditional intelligent tutoring systems (Woelfel et al, 2010), and help children understand the learner model better (Zapata-Rivera, 2003).

Most research investigations in this domain use a fixed set of emotions, usually evolving around Ekman & Friesen's (Ekman & Friesen, 1978) six universal emotions. However, no explicit model of emotions for use by pedagogical agents in ITS can be found in the literature. The literature lacks particularly in explanations and theories with regards to the impact of emotional feedback on user's actions in a pedagogical context. Indeed, most systems seem to use an add-hoc order for the emotions, or make a random selection, like in (Irroyo et al, 2010). This introduces the need for further investigations in affective models for embedded digital characters in educational applications.

A.M.E.R., an Affective Model of EPA's Responses to user-interaction with mathematical drill-and-practice OLM applications, has been introduced in chapter 2. This chapter is concerned with evaluating the different aspects and impacts of the model for OLM applications: Section 3.2 describes the participatory-design process of designing the affective components to be embedded in the *DividingQuest* in order to suit the model's needs, performed with children as design partners. The goal of this PhD's work is not to provide the research community with a new theory on emotions for learning, but to use the existing theories and test their impact on OLM system use when affect is embedded in EPAs. Consequently, section 3.3 will not offer a complete evaluation of the proposed model, in order to introduce a new theory, but merely test the usability of one major decision taken in the definition of the model: separating the feedback according to teacher's expectations for a child concerning a skill level.

3.2 EMBEDDING THE MODEL IN THE *DIVIDINGQUEST*

This section features the participatory-design process conducted, which included two classes of children from French primary schools design partners (17 children aged 7-9; and 27 children aged 9-11). The goal of the study was to produce a set of graphical animations portraying the *DividingQuest* Embodied-Pedagogical Agent, the witch Sélénia (Girard, 2006), in various emotional states to be later used as emotional feedback in the application. As previously discussed in chapter 2, French primary schools are now more and more populated by children from European nationalities, with a strong influence of English families settling in the French countryside. Ten children (four aged 7-9 and six aged 9-11) involved in the process were of English nationality and had only spent a couple of years in France at the time of the experiment. The use of both French and English children, and from two age-groups, should enable the building of animations recognizable by both children populations, without any cultural or developmental differences due to the users participating in the design process.

3.2.1 TECHNIQUES USED AND RATIONALE

The design process involved commonly used participatory-design techniques such as Bekker et al's KidReporter's (2002) user-requirement gathering technique, the Bluebells process (Kelly et al, 2006), as well as results from research on Interaction Design for Children. In this method, two elements were added to traditional design techniques: the inclusion of actors as design partners to help in the expression of emotional display; and the use and creation of comic strips to express rational thought processes.

In a previous study on the design of EPAs for children's educational applications, presented in the Appendix C, comic strips and *mangas* literature were found useful by children aged seven to eleven, and an inspiration for the design of graphical embodied pedagogical agents. Comics were used here as graphical reference for emotional expressions, and a mean of communicating in the process by sketching the ideas of the design partners to show others.

As discussed in the literature review, actors are a great source of knowledge when trying to represent emotional expressions. The actor partners worked closely with children on the different ways to portray emotions on a human body, and its applications to digital characters.

3.2.2 DESIGN PROCEDURE

The study aimed at producing a set of graphical animations including affect, that were usable and recognizable by children aged seven to eleven years. It took place within the context of the “initiation to drama” project taken by both classes involved: children in classes are divided into groups to follow several manual activities (such as life drawing, pottery, theatre, etc...), the activities changing for the groups throughout the year.

In the first step, the teachers from the two classes worked with their students on the verbal and contextual meaning of emotions.

In the second step, children were then divided into groups of six or seven, mixed in age and nationality, and assigned five emotions to graphically represent. They first engaged in a written production activity to describe a specific emotional situation. They then used Ekman and Friesen’s FACS (1978) system to draw the face of the emotional character, before working with their assigned actor design partner on body posture and expression.

In the third step, children were given an illustration of the Traffic-Light metaphor and how it could be applied in the design of a smiley to show affect better. They were asked to use colours, and any accessory they deemed useful to help in understanding the emotion represented better. In order to perform this task, they were given examples of comics, and participated in a class outing to the International Comics Festival of Angoulême (France). They met current and prospective comic designers, authors, and illustrators. They also participated to children’s workshops on how to create comics animations.

Before the next step, designers assembled and incorporated the data into the design of a first version of the final animations. Finally, an iterative process was engaged in which each group of children was asked to recognize a sample of emotions they had not worked on before, with a redesign of the animations in between each sample rotation.

3.2.3 DESIGN PROCESS:

STEP 1: DEFINITION OF THE EMOTIONAL TERMS

The two teachers first engaged in class pedagogical activities to help children understand each emotional state of the A.M.E.R. model: amused; awed ; bored ; comforting ; disappointed ; empathic ; hopeful ; inspired; interested; sad ; satisfied ; unpleasant surprise. The emotions were put in the context of every-day life situations as a first description. Then, children were asked to perform different tasks: recognize how characters felt at a specific moment of a story; watch a short cartoon (with no spoken dialogue) and point out emotional display; or perform activities of “link the concepts” by associating each emotion listed with the real-life example provided. Children were then invited to ask for more clarification when they did not truly understand an emotional concept.

STEP 2: DESIGNING EMOTIONAL CHARACTERISTICS IN DIGITAL CHARACTERS.

Asking each child to design the thirteen animations would have been rich in terms of data, but subject to too many individual preferences and not achievable in the time allocated for the class project. Due to the cognitive capacity limitations of the children who participated in this study when confronted to multiple information at once (Markopoulos et al, 2008), the high number of animations would have been too difficult for children to consider as a whole. Children were therefore assigned to grouped (of six to seven children) mixed in age and nationality, and given a sample of five emotional states to represent graphically. The emotional states were distributed among the groups so that each emotional concept was studied by two to three groups. The states considered by three groups were the ones deemed most difficult to represent graphically from results in the literature: in awe, hopeful, and empathic, as illustrated in Table III.1:

Table III.1: Emotions assigned to each design group

Group	Emotion1	Emotion2	Emotion3	Emotion4
1	Awed	Bored	Comforting	Satisfied
2	Hopeful	Amused	Interested	Unpleasantly surprised
3	Awed	Hopeful	Disappointed	Satisfied
4	Hopeful	Bored	Inspired	Satisfied
5	Awed	Empathic	Interested	Sad
6	Empathic	Comforting	Inspired	Sad
7	Empathic	Amused	Disappointed	Unpleasantly surprised

If the children design partners were to be left alone to perform the different tasks in groups, the participants might not be able to write what happens clearly or quickly enough, or be influenced by outsiders, forget to record some of the data/discussion, or lose focus on what has to be achieved. Taking into account the diversity of reading, writing and oral expression skills of all the children involved, a study coordinator was part of the design process for each group (all adults, unknown to the children prior to the test).

Step 2.1: What do emotions represent?

The children first engaged in a reflective process on the meaning of the five emotional states they were assigned, and how they could be applied to learning strategies. The children worked on the production of emotional representations during the hours allocated to the project, using the expression method they preferred.

In order to suit children's individual preferences in expression style, the learning situations to be defined were allowed in various format: interview with the expert in class, written description of the situation and/or drawings for homework.

- *Interviews*: some children are more comfortable with talking to a stranger about things they did than others (Markopoulos et al, 2008). For such children, the interview technique undertaken by the study coordinator, helped gain some insight on how children understood and mentally represented the emotions. During these group sessions, the other children of the group were allowed to make comments, adding information about their own way of representing this emotional state, or arguing about the emotion in itself. They took turns speaking, and were not compelled to interact if nothing came to mind, or they wished to use another method to express themselves.
- *Written summary*: This is the preferred means of expression for children with high reading/writing skills who are reluctant to speak in public (Bekker et al, 2002). They were asked to describe the emotional states using words and/or pictures, using at least twenty words but not exceeding five paragraphs of ten to fifteen words.
- *Drawing/comic strip*: Children were offered the possibility to express themselves without the use of words (spoken or written), to enable full participation of lower reading and writing skills students (Markopoulos et al, 2008). Once the drawing was performed, they explained it to the study coordinator in one-to-one interaction.

Step 2.2: Designing emotional representations

Once in groups, each of the five emotional states to represent were considered separately, in an iterative process: for each state, the children first assembled the facial expression using FACS (Ekman and Friesen, 1978), as introduced in the literature review in section 1.3.3, and discussed the results to arrive at a final expression; then worked with the actors-partners on posture and attitudes concerning this emotional state; and finally, defined the different steps of the animated representation for this emotion using cardboard body-parts.

In a first step, children were given a drawing of the witch character for which to express affect, void in terms of facial elements. They were also given different forms of facial elements to help represent the face, considering the following facial action units: mouth, nose, eye, eyebrows. Those FACS elements were designed in black thick lines on a transparent sheet of paper, so that the underlying elements could be seen-through. The children, by associating the different elements, and trying them out on the basic face, produced a complete representation of the character's face, as illustrated in Figure III.1:



Figure III.1: Using FACS to design faces.

Once each child assembled a face, the results were looked at by each member of the group, and discussed to form a single model of the emotional state on a separate representation. Once a consensus was reached, the different elements were taped together and a photocopy of the final expression was given to each child of the group.

The following step included a period of theatrical work with the actors-partners, working on ways to theatrically represent the emotional state in terms of posture, expression and attitudes. Different situations in which the emotional state could appear were played out, with the actors acting

out examples of specific postures and giving advice to children on how to position themselves to inhabit this particular emotional state.

In a final step, the children returned to the face undertaken in the first step. They were given a cardboard representation of the character's body, where all the major articulations of the body could rotate (knee, hip, ankles, wrist...). They worked together to rotate the body parts in the "right" position, in such a way that emphasized the emotional state considered. Each expression was then photographed with a digital camera for safe-keeping, which were used again at the end of step 3.

STEP 3: MAKING THE CHARACTERS ALIVE

In this part of the project, children were initiated to different techniques in creating comics and digital characters that express emotions: use of colours and sizes, examples of comics design, and discussion with comics designers/authors/illustrators. The children were given notebooks, considered by some as "project diaries", where they drew or wrote about the emotional characteristics and techniques of design they saw, heard of, or experimented during the whole process (e.g. an image or animation that really moved them, the position of a graphical element that made them think of an emotion spontaneously, ...).

First, the use of colours and size was approached, with a presentation of the Traffic-Light metaphor, and its application in the design of an emotive smiley-character; and a visual presentation of expressions portrayed in different sizes with their impact on the targeted users (what happens when the head of the character is bigger than the average human? ...) Figure III.2 provides an illustration of a smiley-o-meter using the Traffic-Light metaphor:



III.2: Pedagogical Smiley-o-meter

Then, children were given a collection of books on how to draw expressive faces (McCloud, 1993; 8fish, 2008), illustrations of human faces (Ekman, 2007; Simon, 2008), and different styles of comics: asian mangas, MarvelTM comics, French comics (CédricTM, ToupetTM, TintinTM, AsterixTM,...). They perused through the graphical representations given and took notes on the elements they thought 'made the emotion', took notes on the accessories used to reinforce the emotional states, and made notice of a particular representation they favoured.

Finally, the two classes participated in an outing at the International Festival of Comics, in Angoulême (France). After this period of research and experimentation with the techniques and material, children found themselves in the groups formed for step 2. The photocopies of finalized products from this step were reintroduced. Children then added colours, changes in size, and accessories to the characters to emphasize each emotional state. They either “animated” the character by drawing several attitudes, later photographed in succession to see the progression of the expression; or provided a written description on the movements to associate to the static pictures.

STEP 4: TESTING THE EXPRESSIONS’ BELIEVABILITY

From the graphical work provided by all children for each emotional state, designers produced a first version of the animated agents, in Flash animation formats.

Each group of children was given a sample of five animations representing expressions they did not previously worked on. The children had to watch each animation separately, and then identify the emotional state represented. They then watched again the animation and discussed in the group how it could better represent the emotional state. Working with the study coordinator, they produced a set of “advices” on how to ameliorate the graphics, which served as input to the next iterative loop for testing, with a redesign of the animations. This process was repeated until each child had seen an animation of each emotional state.

3.2.4 SPECIFICS OF THE METHOD’S DESIGN TO MAXIMIZE CHILDREN’S INVOLVEMENT

As perceived by many research teams working with children, this group of users can sometimes have trouble providing answers when only relying on verbal skills. Individual liking of a preferred expression method also has a role in the willingness of children to express themselves in one way or another. Children are simply more at ease with interlocutors that speak “their own language”, using a vocabulary and phrasing they are used to (Markopoulos et al, 2008). The variety of methods introduced to children to express themselves with will hopefully limit the number of children staying away from the activity due to a personal limitation, not related to the activity at hand.

In order to enable children to create more design options, supporting children’s multiple intelligences (Gardner, 1983) has been recommended during the design process (Guha et al, 2004; Mazzone et al, 2008; Read et al, 2005). For this method, children use linguistic, visual-spatial, bodily-kinaesthetic and naturalistic skills, in a variety of design activities: low-tech prototyping known to help children to discuss and develop their ideas in a concrete way (Druin, 1999); group brainstorming to encourage verbal expression; and drawing to offer expression methods that suit each child with considerations to their literacy and verbal expression skills (Greig & Taylor, 1999). Children were

encouraged to sketch their ideas during the whole design process in their diaries, as this technique was found to help the discovery and communication of design ideas, by providing reflective feedback (Tohidi et al, 2006).

When in groups, children were joined by a study coordinator, in charge of reminding children of the activity's purpose and process, and mediate the verbal communications to include all participants. From the results of Druin et al (2007), the study coordinator did not take notes during the experiment. A laptop integrated webcam and audio-recorders were used to record children's interaction during the design activities and communications. The apparatus was hidden from view by the children, as it was shown to distract child-users or make them uncomfortable (Druin et al, 1997). Authorization to record the sessions was however required by the child's teachers prior to the design process.

As in the KidReporter's (Bekker et al, 2002) experiments, children felt more comfortable providing information about situations that happened to them or peers in a familiar environment (school, or home with the homework assignment), which helps them form a more decisive description. After the first session of interviews, they felt more comfortable in sharing experiences with their peers, and engaging in discussions on the different emotions felt or perceived, especially during the sessions on the computer.

As stated by Acuff and Reiher (1997), children can easily be influenced by parents or other strong leaders, which bias data gathering. In order to limit this effect, the groups of discussion were formed with the help of teachers, according to children's temperament and interactions with others. a step was taken in order to limit influences of outsiders (such as parents, or siblings, ..) in the writing and drawing assignments: children selecting such means of expression were asked to perform the task in the classroom. This was undertaken during an appointed time, as a single assignment, while their peers were involved in other a classroom-related activities.

3.2.5 DESIGN RECOMMENDATIONS

From the results of this participatory-design study, as well as teachers, actors, and child design partners interviews after the study, a series of design recommendations was drawn when considering the design of affective expressions in digital characters:

- *Use a diversity of experts as design partners: include comic designers and actors for affective products.*

As previously stated, decades of comics design and theatrical interpretation of human emotions have produced rules and tips for actors and illustrators to convey affect in a more accurate and effective way. The input from current comics designers/illustrators and actors were considered as highly useful by children design partners.

- *Use Facial Action Units to manipulate expressions and emotional concepts*

The FACS system, usually used to recognize emotional expressions from humans in static images (Ekman, 2007) was here integrated in an activity to design affect. Children were keen on testing the different positions to form the expressions of emotional states. The use of transparent Facial Action Units helped in the final affective representation, with children moving the elements back and forth on the face to test the visual display.

- *Use objects to help put the emotions in context*

Research in affective computing is mainly focussed on the reading of facial expressions, for which various theories have been drawn. However, a previous study with young children has shown that children understand emotional aspects better when objects or colours are added to the design (cf. chapter 2, design of *Sorémo*). This design process also highlighted children's preference for designing emotional expressions when the context of the emotion elicitation is explicit: they designed a selection of tools for the character to have, and used them to focus the product design on the valence or intensity of the emotion expressed.

- *Vary the activities to help design the whole emotional space*

One of the main issues in using children as design partners is to keep them interested throughout the process. Read et al (2005) advocates that the design activities should include some element of fun, in order to help produce better products. The definition of the design study in several steps, each requesting a different kind of children's intelligence (linguistic, visio-spatial, bodily-kinaesthetic) and ability (literacy, artistic, communication skills), helped keep the children interested

in the different activities, and avoid repetitive tasks that could lead to boredom (Fontijn & Hoonhout, 2007).

Similarly to the results in (Bekker et al, 2002), the definition of the design study in several steps, each requesting a different kind of children's intelligence (linguistic, visio-spatial, bodily-kinaesthetic) and ability (literacy, artistic, communication skills), helped keep the children interested in the different activities, and avoid repetitive tasks that could lead to boredom. It also had another impact on the study design: the use of different means of expression helped the identification of the different elements to be included in the expression of affect. Some children found it easier to describe the position of the agent's body or facial expression through writing, drawing, or oral expression. It also facilitated the creation of objects and to put the emotions in context for better understanding of their role in future software by children.

- *Integrate the design study in a class project*

Another issue in using primary school teachers and children in the design process is for the design partners to engage in creating a product they might not use once finished. Indeed, due to the length of time needed for the design process, and then the implementation of the product, a school year is generally over before the end of the whole process. By integrating the activity in a class project, with a concrete output (here the definition of a small theatrical representation on the expression of emotions, using the cardboard characters on a giant background display), children and teachers see the finality of the activity, and are more willing to see the project throughout. The teachers can then reuse the material created in the process to carry on working on a pedagogical activity, and present the results of the work to other teachers or parents.

Once the animations were designed and validated in terms of usability, they were integrated into the *DividingQuest* learning environment, according to the feedback defined in the A.M.E.R. model introduced in chapter 2. This method appears promising in terms of designing affecting components for educational software with children as design partners. One can wonder whether the design object – EPAs to be embedded in a pedagogical game – could have impacted how well the result fit with the choice of adult design partners – comic designers and actors –, and whether this can be extended to other uses of affective design. The results of the application of this method for the construction of the EPAs seem encouraging on that aspect. It is true that the use of comic designers as resource for children's design of emotional display suited very well a gaming and fun concept. However, the children were not involved in the process of building the educational application. The animations, product of the design, were introduced to the children as a tool to be used in a pedagogical context, via the use of the scenarios of emotional display. However, they were not presented as part of a specific application or for a specific pedagogical scenario.

Children found the use of acting as a method to inform their design, and the wealth of knowledge the actors shared, essential to help them put the emotions in context and visualize how to different emotional representations change how humans react to the emotion itself. The main drawback of this method is its cost with the use of actors, comic designers, teachers, and children in the process. However, the cost of design might be reduced by integrating the knowledge of such participants into the design team. Actually, most design teams have a number of people involved from different disciplines: teachers, academics in different domains of expertise, designers,...The next section evaluates the impact of the affective responses chosen within such a model on children's use of the *DividingQuest* OLM application.

3.3 TESTING THE MODEL'S VALIDITY

3.3.1 AIM OF THE PILOT STUDY AND PROCEDURE

In order to validate some of the design choices made in building the affective model of EPA's responses previously presented, a pilot study was conducted on twenty-seven children from an English primary-school, and twenty-seven children from a French primary-school, aged nine to eleven years. It aimed at analyzing children's emotional reaction to such feedback and the impact on learning of the various emotional strategies involved.

It aimed at answering the following research question:

- Does the choice of emotional response in EPA impact the child's emotional responses as a function of their knowledge level prior to the activity? (e.g. do children react differently according to the emotional strategy followed?)
- Does the value of the emotional feedback as defined in the model in terms of engagement level and motivation differ as a function of different mathematical ability?

The pilot study consisted of a learning session of thirty minutes on the respective software (*DividingQuest* for the English children, *Multipliotest* for the French children). Children's interaction with the software were recorded, as well as their emotional reactions to software interaction via webcam recordings, built into the computers used for the experimentation. The children knew the session was recorded, and written consent from parents was given for each participants. However, the position of the webcam ensured a low level of distraction for the child users concerning this piece of knowledge.

3.3.2 STUDY DESIGN

The study was designed as a 3 group between-subject design, with each child working on one unique user-interface. The goal of the study is to validate the benefits of separating of feedback according to children's prior knowledge for better learning and to keep children motivated. For this reason, each type of feedback was tested by children from all knowledge groups, in order to identify which feedback seemed better suited for each group. The two groups of twenty-seven children were each represented by nine children of each level (low, medium and high) of the mathematical skills involved in the software. Children in the English primary school worked on the concept of divisions, using one interface of the *DividingQuest*. Similarly, children in the French primary school manipulated the concept of multiplications, by using one interface of *Multipliotest*. Each interface was seen and manipulated by nine children, of mixed abilities (three of each level). The groups were divided using their level's expertise on their mathematical skills abilities.

Table III.2: Repartition of children per interface

Skill Level Interface	Low(9)		Medium (9)		High (9)	
	DQ	M	DQ	M	DQ	M
Able	3	3	3	3	3	3
Borderline	3	3	3	3	3	3
Unable	3	3	3	3	3	3

Each interface illustrates one strategy of emotional responses to EPAs, as described in the model in table III.2: the “able” interface includes emotional strategies initially designed for the more skilled students, while the “borderline” interface features emotional responses designed for medium-skilled children, and the “unable” interface for poorly-skilled students.

The children considered with “low” division skills were divided into three groups, each group being submitted to one only of the Affective Model simulation of EPA's responses: one where the student was considered able to do the activity, one for borderline students, and one where the student was not considered able to perform the tasks correctly. The same distribution of interface per child occurred in all the groups aforementioned.

3.3.3 RESULTS

Three types of information are considered within this study: the impact of an EPA's emotive intervention on children's learning strategy, help-seeking behaviour, and the user's emotional reaction. The emotional reactions were singled out from the video stream recorded, and compared with Ekman & Friesen's FACS (Ekman & Friesen, 1978) to identify the underlying emotions. When the children were interviewed, they were asked to corroborate the identification of emotions by looking at the video recording of their session and to explain their reaction at this particular moment of interaction.

Several different types of behaviours could be observed, and classified according to Roll et al's meta-cognitive ACT-R model of students' learning strategies (Roll et al, 2004). A pilot study was performed with three children, one child working on each interface of the *Multipliotest* software. After obtaining consent from their parents, a video recording was made of their performance on software. Three researchers then reviewed the video recordings and compared children's behaviours to the information recorded by interaction logs. A coding system was developed from these data sources, inspired from Roll et al's learning strategies (Roll et al, 2004). Once each child performed a session on software for the study, his/her behaviour was coded using this scheme by two independent researchers. The results were compared, and items that were found in disagreement were considered again by all researchers. A discussion took place debating on the final coding scheme for said items, until an agreement was reached. This qualitative approach to study validity, or reliability, was chosen due to the nature of the study, and the number of participants involved. It indeed facilitated the resolution of conflicts in experts' judgment by negotiation.

One behaviour is looked at here in particular, corresponding in terms of student's goal as "Least Effort Oriented": the child begins an activity, tries a few questions, and then does not attempt to finish it, but tries other activities instead, one after the other.

Two types of children appear to followed this behaviour when confronted by pedagogical software: low ability students who find the activities too difficult and "give up"; and high ability students who find the activities too easy and "flutter" from one activity to another in order to find one more interesting or challenging. More behaviour of this sort was found in the low and moderate students working on the 'able' interface than the high level students, or the low and moderate students on other interfaces. When interviewed on the subject, children stated that the lack of understanding and empathy of the EPA when they failed to give the right answer discouraged them from trying the activity further. A second tendency has been revealed for the presence of this behaviour: highly skilled students quit activities abruptly more in the unable interface than in the others, when they tried the "easiest" activities to try on. This suggests that the type of feedback given to low-level students in the model might lead highly skilled students to work less hard on software and explore/play more than

when confronted to the feedback from the model. As children were free to choose the activities they wished to work on, not enough data was recorded in this pilot study to see the impact on learning. Impact on learning, however, was not at the centre of this study and will be investigated later in chapter 6.

A greater level of frustration was identified in the “able” and “unable” interfaces. Respectively, more lower and higher level students were found to express frustration during the game activities than the other children working on this interface. On interviewing the children, it appears that the high level (66%) and some moderate level students (33%) working on the low level interface found the emotional display to be too compliant and too comforting, lacking in challenge, especially in the easiest activities of the learning platforms. On the other hand, the low-level children working on the most challenging interface (100%) seemed to have trouble keeping up, and felt discouraged by the character’s display, whereas children of low level working on the borderline and unable interfaces expressed a low level of frustration and appeared more motivated to keep working on an activity until they understood its principles.

From these observations, the design choices concerning the type of feedback to give children according to their prior level seems appropriate. Indeed children under the feedback initially designed in the model seemed less frustrated using software, and less activities were quitted suddenly after only a few tries.

3.3.4 CONCLUSION

A number of observations have been made from the study results:

- Children's behaviour towards choosing and keeping a level of engagement with an activity when confronted with a particular EPA's emotional response strategy, varies according to the child's knowledge prior to the activity.

- Children's emotional responses to software interaction when confronted with a particular EPA's emotional response strategy varies according to their knowledge level prior to the activity.

The differentiation of EPA's emotional response to user interaction according to his/her acquisition of skill level can provide users with adaptive learning. The children understand the emotions chosen and their representations, and this impacts the decisions they made as to software interaction. It appears that the choice in responses, when the interface tested corresponded to the child's level, helped to keep children more motivated, and avoid jumping from one unachieved activity to another, when the appropriate emotional strategy is used. Some children from the low level group stated that after working on the unable or borderline interfaces, felt that the EPA had taught them something, and felt more secure in answering mathematical problems.

The pilot study is a first step in validating the use of different emotional strategies according to the teacher's expectations concerning the child-users, defined by their current level of skills concerning the application pedagogical content. The next section will now investigate children's emotional responses to the A.M.E.R. feedback.

3.4 SUMMARY

This chapter was concerned with the content, design, and link to the pedagogical activities, of the affective components of OLM intelligent tutoring systems used by French and English primary school children aged seven to eleven. It aimed at answering the following research questions:

RQ1: What theories of emotion can be followed to produce a model of EPA's affective response in OLM software suitable for mathematical drill-and-practice applications?

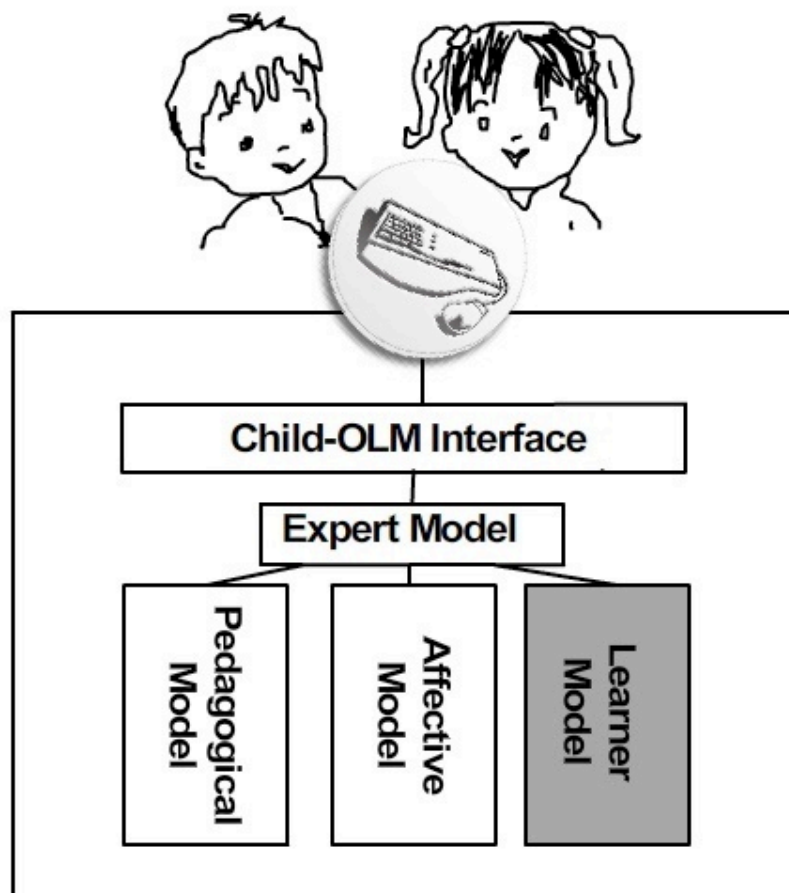
RQ2: How can user-centred and participatory design techniques be used with children and their teachers to create affective components to be integrated in tutoring systems for children?

The pilot study presented in section 3.3 indicates that the use of different emotional strategies according to the teacher's expectations concerning the child-users, defined by their current level of skills concerning the application's pedagogical content, is viable. This justifies the use of the A.M.E.R. model for OLM drill-and-practice applications aimed at young users aged seven to eleven in this PhD's work. The model has been included in an investigation as to how children respond emotionally to the A.M.E.R. model. The results showed that on average, children responded similarly to the participatory design's predictions to the EPA's feedback. It will be tested in terms of benefits for learning and motivation when compared to other types of feedback to child-interaction with the environment (see the experiment presented in chapter 5).

The second output of this chapter concerns the use of participatory-design techniques with children to build affective components and their teachers to adapt theory and school practices for technological application and utility. Using children as design partners to design representations of affect understandable by this user age-group through the design of affective digital character animations, helped reach an acceptable level of emotion recognition for such animations (average of 92% recognition). The use of comics and mangas literature, as well as role-play in the participatory-design process was deemed essential to the success of the design activity. Practical experience concerning the design process led to the production of some design recommendations.

This concludes the discussions about the design of the affective components of OLM mathematical software for young children. Before exploring the benefits of opening the model to young users in terms of learning gains and motivational gains, chapter 4 will investigate how the learner model can be visually represented in OLM software for children to reach a better user understanding of the information provided, and their possible use for better learning.

Chapter 4: Designing the content of the learner model



4.1 INTRODUCTION

The learner model of an Intelligent Tutoring system is often defined as “a student-instantiated version [referred to as “view” in this PhD] of the competency model. That is, values in the student model express the assessor’s current belief about a student’s level on variables within the competency model” (Shute & Zapata-Riviera, 2010). The views of such OLM can vary in terms of content – from a simple overview of knowledge level to interrelationships between concepts and/or misconceptions – and representations – from basic Likert scales to Bayesian representations in 3D graphs. Each learner does not integrate and visualize knowledge in the same way (Bull & Kay, 2010). Therefore users do not reach the same level of understanding of the information the LM provides according to the way the LM is designed and displayed, and thus they react to the display accordingly as regards to motivation of use.

This chapter is concerned first with the study of how children’s state of knowledge can be represented for children to understand, and second with the modelling of learner model components in such a way as to invite child-users to interact with them. An exploratory study is presented evaluating how children aged seven to eleven years perceive and use different visual representation tools and LM views within OLM computer-assisted learning systems.

4.2 WHAT REPRESENTATIONS OF LEARNER MODELS DO CHILDREN FAVOUR AND BEST UNDERSTAND?

A learner model view is composed of a set of LM tools, organized in a specific way. This study is investigating how the content of the model can be represented in a way most appropriate for children, in order to maximize their understanding of the content, and by taking into account their general visual preference.

This study aims to answer two research questions:

- Which modelling tools, commonly used to represent LM content in the literature on OLM systems, are best understood, and preferred, by children aged seven to eleven years old?
- Which modelling tools, organized under which view, are more likely to be understood and used in OLM software for children this age group?

In order to answer these research questions, child-participants were involved in two studies of the use of modelling tools (Study IV.1), and their association to LM views (Study IV.2). children’s

preferences in terms of tools and views, and their understanding of the underlying LM content, were quantified and analysed using descriptive statistics.

4.3 LM VIEWS AND REPRESENTATION TOOLS.

4.3.1 ITEMS OF LEARNER MODEL CONTENT: LM REPRESENTATION TOOLS

Six LM representation tools (Skill-Meter, 2D-Graph, Boxes, Tables, Text, and Bayesian Graph) were chosen for this study. They represent the most commonly used visualisation tools in OLM software modelling (Bull & Kay, 2010), as described in the literature review in section 1.4.3.

Figure IV.1 represents an example of each tool for one item of learner model content:

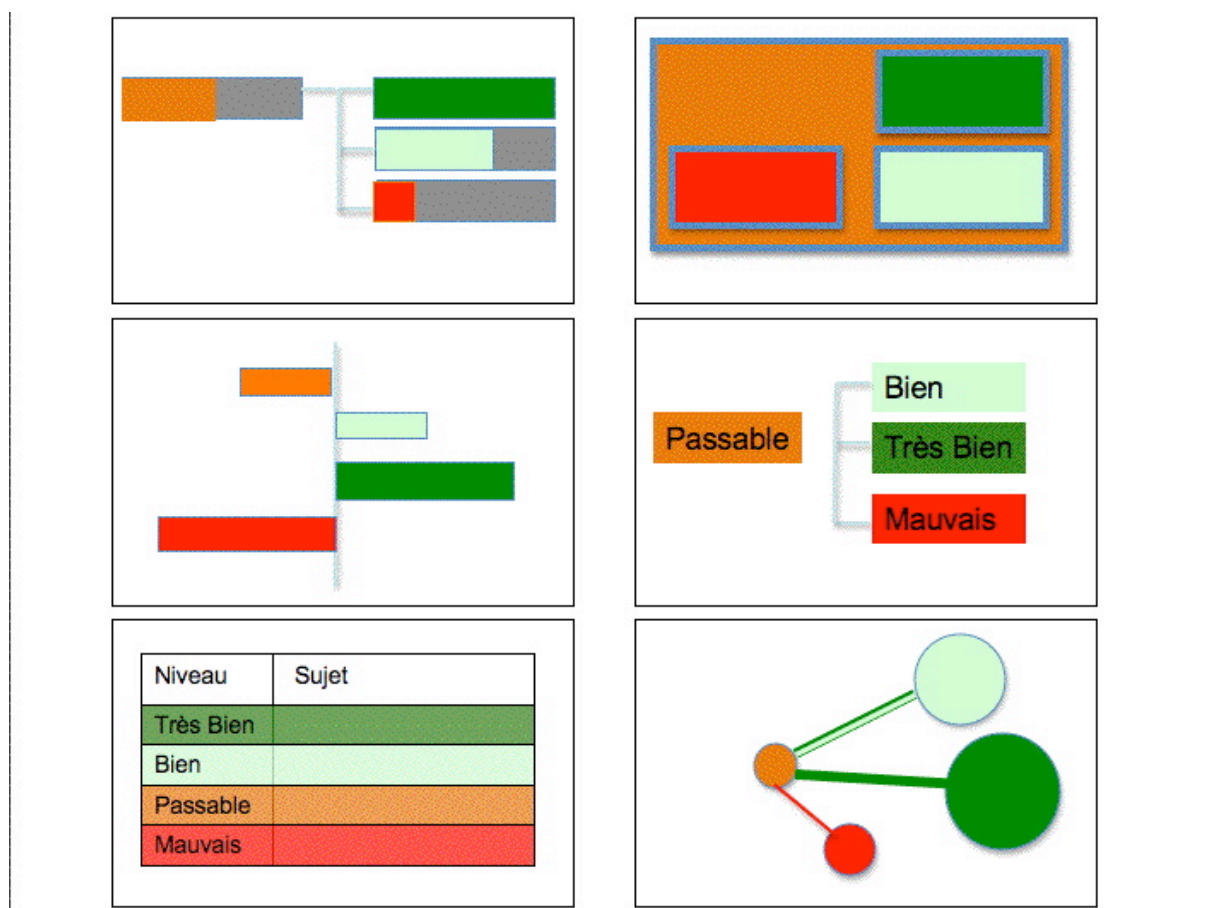


Figure IV.1: The six LM tools used
(Skill-Meter, 2D-Graph, Boxes, Tables, Text, and Bayesian Graph)

In study IV.1, the learner model view was restricted to the ‘related concepts views’ as defined in (Mabbott & Bull, 2004). Each of the six modelling tools was used in the representation of one learner model content, under this view, and integrated into an OLM prototype for viewing by the

children. The representation of the model chosen to help children in their understanding of the underlying concepts is simple in essence: Each tool uses the traffic-light metaphor, which proved to be sufficiently understood by children this age (cf. appendix B) to be viable for use in experiments, and is widely used in the OLM community (Bull & Kay, 2010). The learner's knowledge of a topic is represented by a single colour code – red, orange, and green – with bright green indicating complete knowledge and red indicating none. This simplicity of representation means that learners should require little time to familiarize themselves with the environment.

Figure IV.2 illustrates a model represented within this view, using Bayesian graphs as a modelling tool.

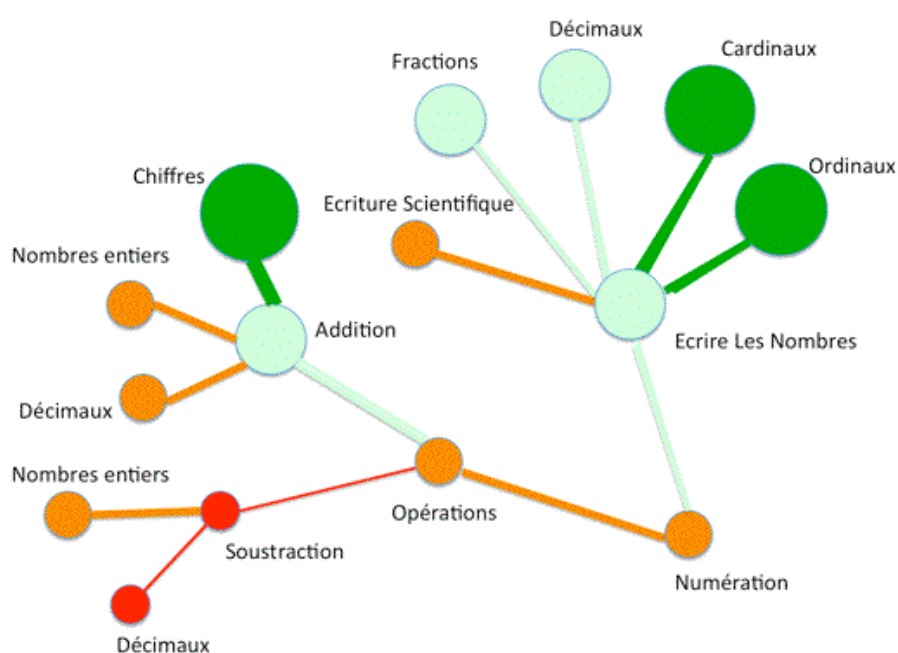


Figure IV.2: Related Concepts view of a learner model using Bayesian Graphs

4.3.2 ORGANIZATION OF THE LM COMPONENTS: LM VIEWS

Learner Model views are a representation of the LM content, defining a specific way for LM information to be linked together on the screen. It can be designed using one or several representation tools. In study IV.2 four views were studied, defined as described in (Mabbott & Bull, 2004):

- lecture view (lists topics according to the order they are presented in the activities of the software),
- related concepts view (logical, hierarchical structured grouping of subject matter),
- concept map view (represent the conceptual relationships between concepts),
- and pre-requisites view (shows a suggested order for study topics).

Figure IV.3 illustrates an example of a learner model, under each of the four views, using skill-meters as a representation tool: lectures view in the top-left corner, concept map view in the top-right corner, related concept view in the bottom left corner, and pre-requisite view in the bottom-right corner (a larger figure of each view is available in the Appendices, section B.2).

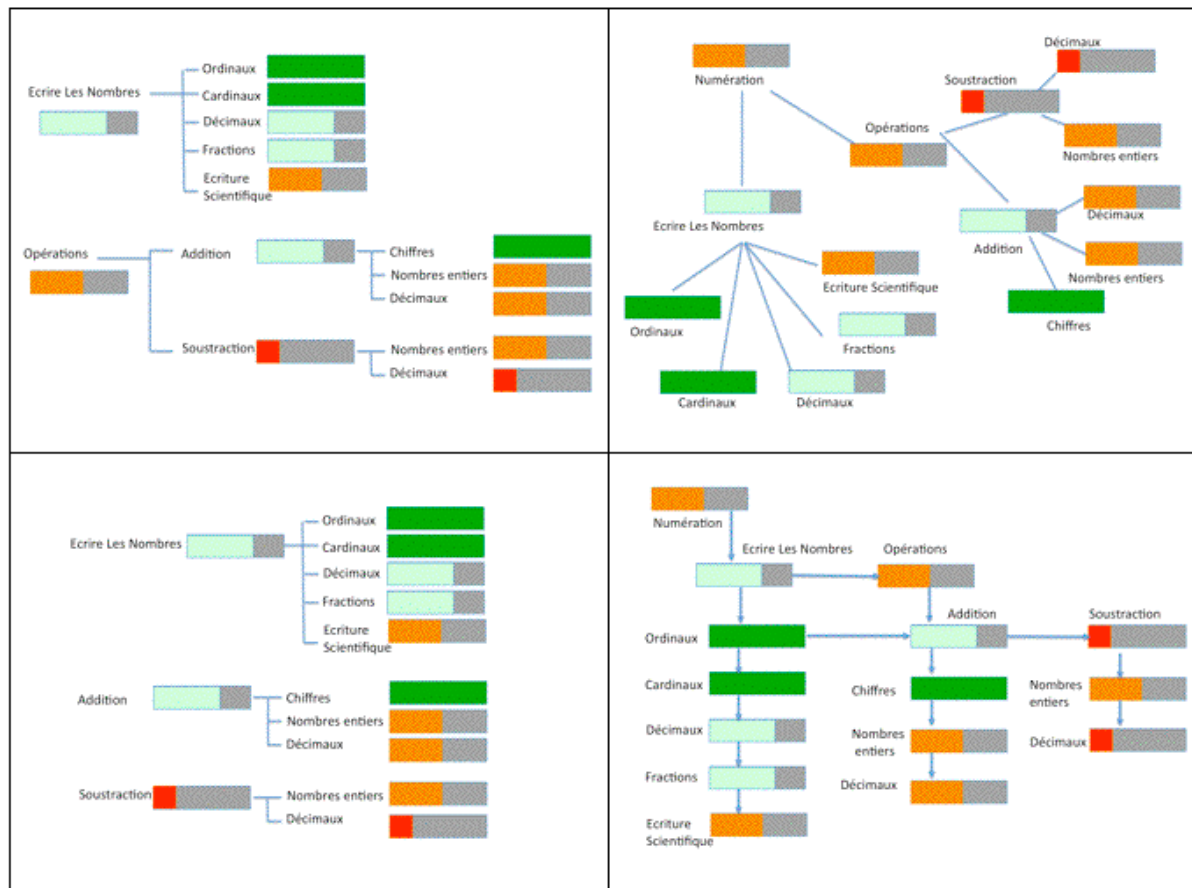


Figure IV.3: Learner Model represented by skill-meters under four LM views.

In Study IV.2, each model used for the experiment was represented under the four views, using each representation tool chosen for the study.

4.4 STUDY DESIGN AND PROCEDURE

4.4.1 PARTICIPANTS

A hundred children aged seven to eleven years old from two French primary schools participated into the study: twenty aged seven, twenty aged eight, twenty aged nine, twenty aged ten, and twenty aged eleven. The children performed each part of the study in a classroom separate from their peers, by groups of 5 participants.

Each child was assigned a study coordinator who guided him/her from his/her classroom to the experiment room, and their assigned computer, where they performed individually. The five study coordinators began the procedure at the same time. The study coordinators recorded all interaction with the child-participants on paper. Upon completion of the study, they engaged in drawing or storytelling activities with the children while waiting for the last of them to finish before accompanying them back to their teacher for the next participants.

4.4.2 RESEARCH QUESTIONS

In this study, the possible relations between children's age, their preference and ability to use LM representation tools, and their view of the LM were investigated:

- Is there a difference in the understanding of the learner models, or preference of use, between the different children's age groups when submitted to different LM representation tools?
- Does the understanding and preference of LM representation tool vary according to the view of the model presented to the user? Is there a difference in understanding and preference in terms of children's age groups?

4.4.3 PROCEDURE

STUDY IV.1

The goal of study IV.1 was to identify which tool or tools facilitated children's understanding of the learner model content, when used to represent the model under a "related concept" view. In order to avoid any bias due to children's recollection of a learner model content from one representation to another, children were provided with six different LM contents to examine. Each of the six screens show to the children was therefore different in content, and used a single LM tool under this view, as illustrated in table IV.1. The 'related concepts' view was chosen for this part of the

study because of its strong resemblance to the progress report given by teachers in this primary school. It was therefore supposed to be the most familiar to children.

Table IV.1: Description of the models used in Studies IV.1 and IV.2

Model	Tool Used	Description
1	Skill-Meter	Mathematics
2	Boxes	English vocabulary
3	Graphs	French grammar
4	Text	History
5	Table	Geography
6	Bayesian	Sport

The models were considered of equal complexity by the participatory-design teachers, and were constituted of the same number of sub-level concepts to acquire. The content was inspired from the school's curriculum, and therefore represented concepts that were common to all children participants, to various extents of knowledge. For this part of the study, each model was integrated into the OLM viewer, a local HTML and Macromedia Flash Application, using one single LM tool, under this one view.

Each child was individually shown the six models, in a random order. The randomization of model viewing should therefore reduce the impact of transfer of model understanding from the use of one LM tool to another within the experiment. For each model, the child had to inspect the model, and then answer a total of five questions on the model comprehension. The procedure was inspired from Mabbott & Bull's experiment (2004). Each model "belonged" to a specific child, as illustrated in Figure IV.4, with 'model 1' on numeracy concepts:

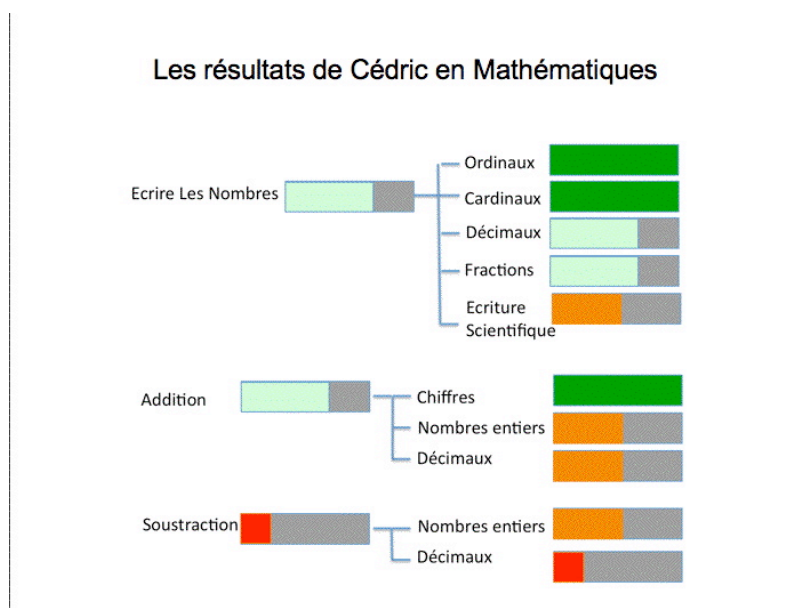


Figure IV.4: 'Model 1' representation with Skill-Meters.

The names were chosen from a pool in such manner as to avoid including the name of any participant. During preliminary work of participatory design and evaluation performed with children this age-group for the PhD, the use of the name of one of the children was indeed found problematic. The peers of said child became confused as to whether the task was considering an imaginary person or the child they knew, and some of them gave answers that took into account the real child's characteristics, rather than the task offered to them.

The following questions were asked (in French, using a wording children were familiar with) for 'model 1', on numeracy:

Q1: Can you describe what is on the screen.. What does it show?

Q2: How good is Cedric at Maths?

Q3: Can you tell me where Cedric does well?

Q4: How good is Cedric at doing additions?

Q5: Does Cedric know how to write numbers well?

The five questions represent different degrees of understanding: general understanding of the information provided, looking for a specific information, understanding a 'node' of items (LM item that is linked to subaltern concepts), ... It also included two types of question generation: closed questions about the level of the child on a subject, and open-ended question looking for a subset of the whole. The later involved the understanding of the notions of seriation, classes, and ordering, which develop around the age of six according to Piaget (1970).

Once the child answered each question for the six models, s/he was asked to rate the representation tools according to his/her preference of use. S/he was presented with cards representing each of the six tools for a simple concept item, as portrayed in Figure IV.5. Each tool was on one individual card, and a color-o-smiley scale was designed on the activity card to complete. S/he had to select each card, and place it in front one of the card spaces, according to how understandable s/he found the tool, from the least (bottom rectangle) to the most (top rectangle) likely to be understood. Figure IV.5 gives an example of child's answer, with the tools ordered from best to worst as: skill-meter, box, graph, table, text, and Bayesian graph.

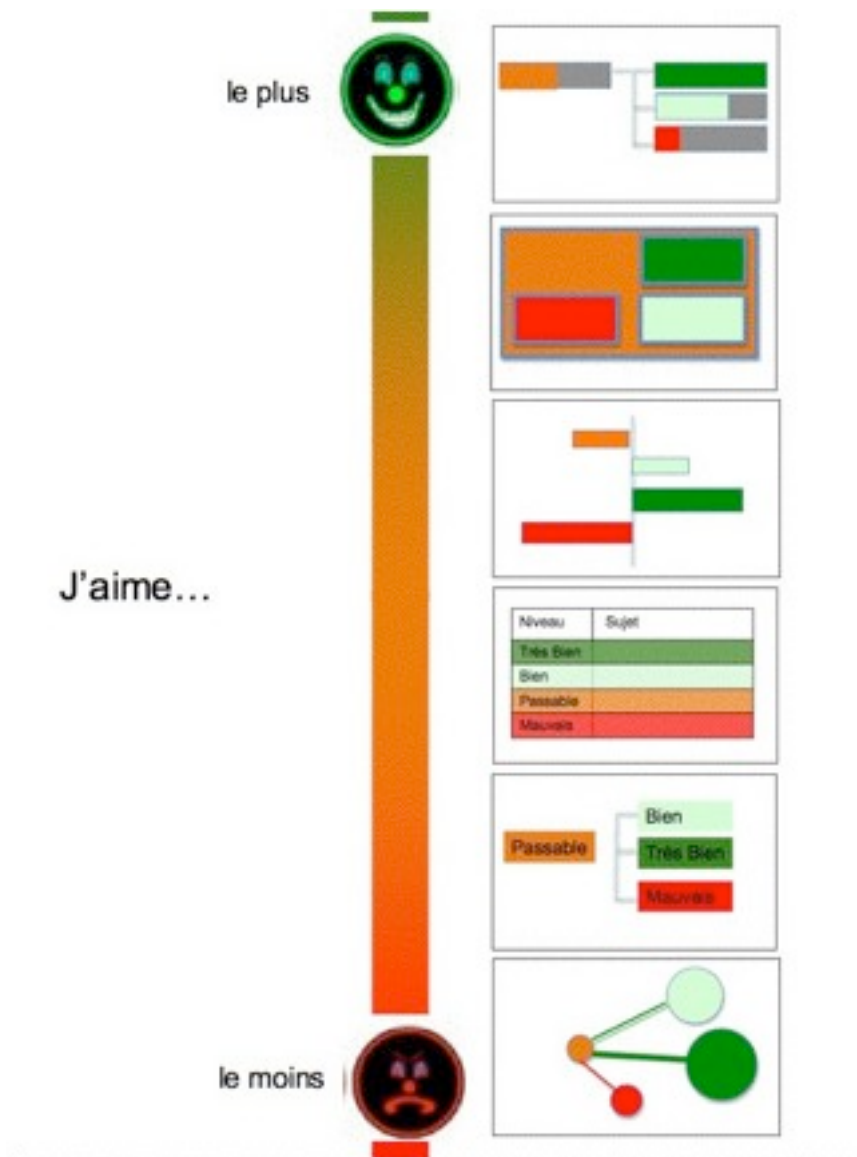


Figure IV.5: Child's preference of LM tools, an example

STUDY IV.2

Study IV.2 took place exactly one week after children's completion of study IV.1. For this study, the main aim was to identify any difference in children's understanding of a learner model representation, when comparing the use of a specific tool, under a several LM views. From the results of study IV.1, it appeared that children had difficulties understanding the Bayesian graphs representation tool, especially for children under the age of nine. This tool was therefore taken out of this study, to concentrate on the other five tools: boxes, 2D-graphs, skill-meter, table, and text. Each of the six models developed for study IV.1 was redesigned for representation by the five tools, under each of the four views : lecture, related concept, concept map, and pre-requisites views. An example of 'model 1', represented with skill-meters under the four views, is featured in Figure IV.3.

The hundred children were divided into five groups of twenty children, and assigned a learner model tool. They were put into groups by their teachers, in the hope to obtain an equal representation between groups of children's age (from seven to eleven), and general schoolwork abilities. Each child was shown four learner models, and asked the five questions corresponding to the model, as defined in study IV.1. Each learner model was represented using the LM tool the group was attributed, under one of the four LM views. The attribution of models under one particular view was randomized, with one child being shown four different models (one for each view).

Once the child answered the questions for each view, s/he was provided with cards representing the four views and a color-o-smiley scale to rate their preference of views. The procedure followed the one described in study IV.1 for the preference of tools.

4.4.4 EXPERIMENTAL MEASURES AND ANALYSIS

Each child performed each part of the study with a study coordinator, who transcribed his/her answers to each question, for each learner model inspected. The coordinators were provided with an answer sheet per model, per child, encoded as follow:

- For Study IV.1: age group, model number, tool number
- For Study IV.2: tool group, model number, view number

Each answer sheet was then analyzed by the three researchers involved in the project: each question answered was marked out of two points, rating the child's understanding of the model inspected for the information requested. A mark was therefore given out of ten for each model viewed. All marks were processed and comparison was made between the grading of each researcher. When a difference greater than two points out of ten was found, the researchers looked at the corresponding sheet together again, and argued their views until they reached an agreement.

Once all models were graded, they were coded using a binary system according to the understanding level: a model was considered as 'well understood' if graded seven or above out of ten. This number was fixed after a meeting between the participatory-design teachers, authors of the model content, and used to marking children's understanding of learning concepts in their everyday activities. However, in order to ensure a fair evaluation of skills, the researchers coded the 'good understanding' of all models with grades between six and seven out of ten by hand.

The results were then analysed considering the number of models that rated a 'good level' of understanding, according to the type of LM tools (for Study IV.1 and IV.2) and views used (for Study IV.2 only), as well as the age of the participants. The 'good level of understanding' therefore represents what the participatory-design teachers and researchers considered as the minimum required

to make use of the learner model information in an OLM, in line with Bull & Mabbott's research (2004).

4.5 RESULTS

4.5.1 STUDY IV.1: USE OF VISUAL REPRESENTATION TOOLS

CHILDREN'S UNDERSTANDING OF THE MODELS FROM THE TOOLS

Table IV.2 illustrates the main result of the study: how many children were able to express their understanding of the learner model content to a level deemed satisfactory for further use in such applications. For each representation tool, a model per child was tested, which brings the maximum number of models "well understood" to twenty per table cell. The results are displayed according to the participants age groups, in order to study any influence of age on the model understanding.

Table IV.2: Models recognized to a "satisfactory" level of understanding, addressed by representation tool used, and age of the participants.

Tool	7 yrs	8 yrs	9 yrs	10yrs	11yrs	TOTAL
Skill meter	15	12	18	19	20	84
2D Graph	2	5	4	10	13	34
Boxes	9	10	18	16	19	72
Tables	1	0	7	7	12	27
Text	0	0	2	1	6	11
Bayesian Graph	0	0	0	0	1	0

Form the data presented in table IV.2, we can see that children did not understand the learner model content to the same level when they were confronted with different learner model representation tools, and that the understanding evolves as children age. For example five children out of the twenty eight-years olds were able to correctly understand the learner models represented by 2D-graphs. On the contrary, all twenty eleven-year-olds appeared to understand the models when presented with skill-meters. The following sections explore the data tendencies in terms of LM tools, and age groups.

General understanding level of the tools:

Figure IV.6, shows the number of models that were correctly explained across all participants age-groups, according to the LM tool they were presented with.

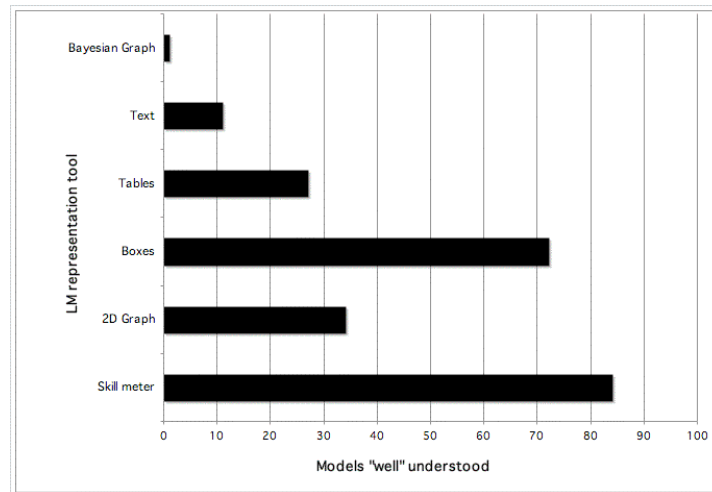


Figure IV.6: Representation of children's understanding of LM visualization tool across all age-groups

It appears that the “Skill Meters” and “Boxes” tools are more likely to be understood as LM visualization tools rather than the “Text” or “Bayesian Graphs” ones.

The “Bayesian Graphs” tool was particularly found hard to understand by all participants, and only 5% of the whole participants succeeded at gaining some understanding, albeit not perfect, of the underlying learner-model.

Impact of subject's age on the understanding level of the tools:

Some of the tools seem to be recognized differently according to the participant's age: While most children aged seven to nine seem to have trouble reading the instructions labelled in text-form, the eldest groups, fully able to read the information provided, expressed their confusion as to the relation between what was written and what they were asked to do, i.e.: explain the learner-model presented.

Two visualization tools are inspected further, as they appear likely to be more easily understood in Table IV.2: the “Skill meter” and “Boxes”, illustrated respectively in Figure IV.7 and IV.8.

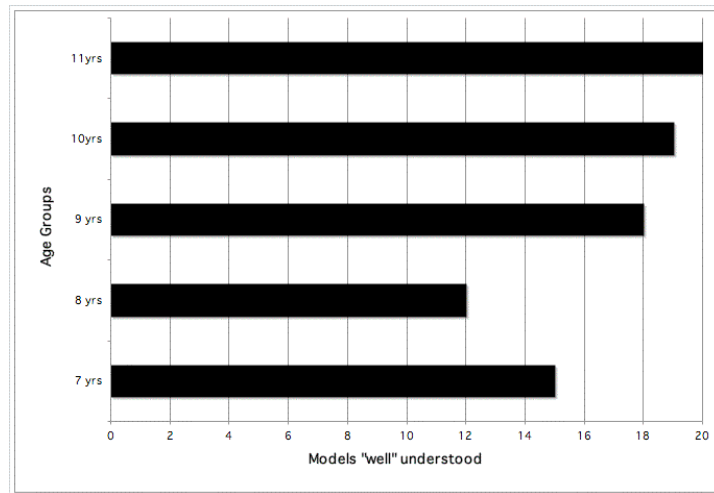


Figure IV.7: Representation of children's understanding of « Skill Meter » by age group.

Inspection of the data seems to indicate that children are more likely to fully understand the concepts represented by the LM when using Skill-Meters as they age. Children aged nine to eleven years seem to have a really good understanding of the system, with more than 72% of the children able to see the underlying notions and how they connect to one another. The population aged eight yrs seemed more hesitant in their understanding, with only 40% gaining a good understanding.

The youngest subjects were found to understand the system quite well, with 90% of them linking well notions and achievements. The remaining 10% (two children) had difficulties in overall schooling subjects, and were the youngest of the class, just having turned seven a couple of weeks before the study. When interviewed, 85% of all children seemed to understand the concept of the Skill-Meter representation. However, for those who did not reach a good level of LM representation understanding, they expressed trouble in relating it to explain the learner-models featured.

Figure IV.8 now illustrates another representation tool, “boxes”, second best recognized by children:

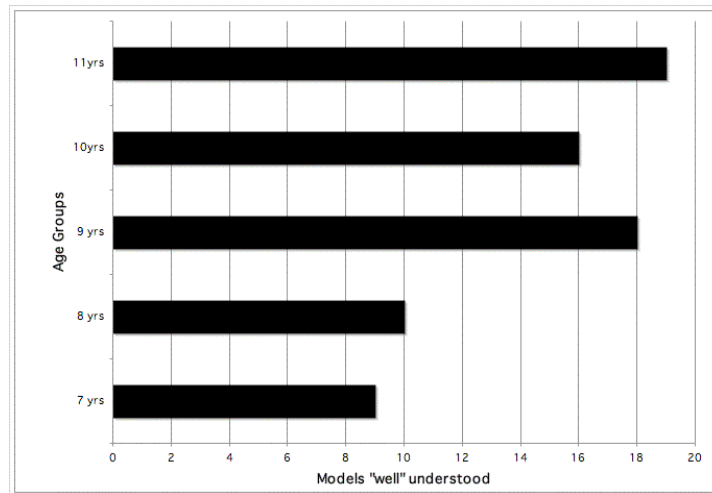


Figure IV.8: Representation of children's understanding of « Boxes » by age group.

Inspection of the data indicates that children are more likely to fully understand the LM represented by Boxes as they age. Children aged nine to eleven years seem to have a really good understanding of the system, with more than 80% of the children understanding it at a good level. The younger populations aged seven and eight yrs seemed more hesitant in their understanding, with only 45% to 50% able to link notions together. Once again, the children having scored poorly in those two groups were the youngest of the group, the two children in the eight years having very low attention span and being easily distracted in all schooling activity (information taken from their teacher).

CHILDREN'S PREFERENCES OF VISUAL REPRESENTATION TOOLS

Figure IV.9 pictures the representation of visual representation tools liked the most (on the left) and least (on the right) by all children participating to the first part of the study.

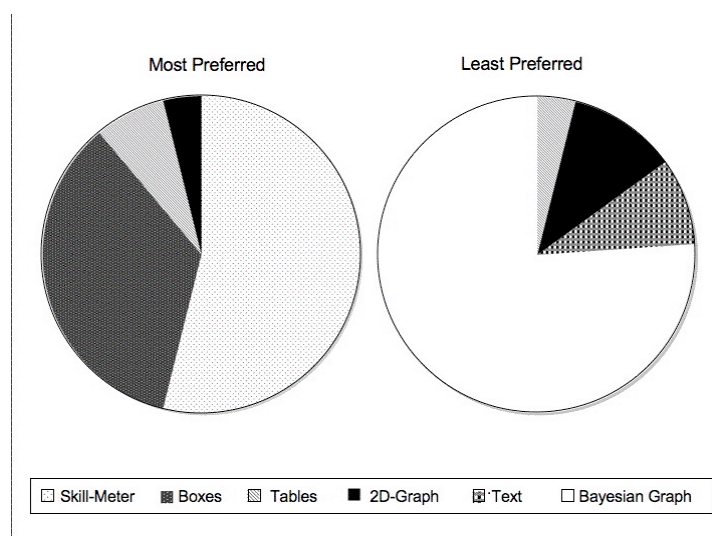


Figure IV.9: Children's preference of visual representation tools

A two-tailed chi-square “goodness to fit” analysis performed on the rating of children’s preference between visual representation tool (respectively most and least preferred) showed a significant difference in the proportion of visual representation tool chosen by children as their most preferred ($\chi^2 = 152.36$, $df = 5$, $p < 0.001$) and least preferred ($\chi^2 = 259.64$, $df = 5$, $p < 0.001$) tool. It appears that a majority of children prefer the “Skill-meter” and “Boxes” visual representation tools, while they least favour the “Bayesian graph” tool, and, to a lesser extent, the “Text” and “2D-Graph” tools.

When applying a score from 6 (most preferred) to 1 (least preferred) to children’s preferences, the following scores were obtained, presented in table IV.3:

Table IV.3: Children’s preferences of LM tools.

Tool	(6) love	(5)	(4)	(3)	(2)	(1) hate
Skill-Meter	54	21	16	8	1	0
Boxes	35	7	43	15	0	0
Tables	7	26	5	13	45	4
2D-Graph	4	16	11	14	44	11
Text	0	16	25	50	0	9
Bayesian Graph	0	14	0	0	10	76

A chi-square ‘goodness to fit’ analysis performed on the LM visual representation tool showed a significant difference in children’s preference of the tool, from what could have been obtained by chance alone ($\chi^2 = 729.96$, $df = 5$, $p < 0.001$).

The results show a tendency from children to order their preference in the following way: skill-meter, boxes, text, tables, 2D-graph, and last Bayesian graphs. The Bayesian graphs, in particular, rated poorly in terms of user preference and understanding, especially for the youngest children.

4.5.2 STUDY IV.2: STRUCTURE OF THE LEARNER MODEL: THE LM VIEWS

CHILDREN'S UNDERSTANDING OF THE MODELS WITHIN THE VIEWS

Four types of LM views were proposed to the children: lecture view (lists topics according to the order they are presented in the activities of the software), related concepts view (logical, hierarchical structured grouping of subject matter), concept map view (represent the conceptual relationships between concepts), and pre-requisites view (shows a suggested order for studying topics). For each view, five visual representation tools were proposed: “Skill-Meter”, “Graph”, “Boxes”, “Tables”, and “Text”, with each child being submitted to one visual representation tool, under the four different LM views (as illustrated in Figure IV.3 with the use of skill-meters to represent ‘model 1’, under the four different views).

Table IV.4 illustrates the number of models that were found as “well understood”, under different views, according to the visualization tool considered:

Table IV.4: Models recognized to a “satisfactory” level of understanding, addressed by representation tool used, and LM view representation.

View	Skill-Meter	Graph	Boxes	Tables	Text	TOTAL
Lectures	20	16	18	3	8	77
Concept map	4	0	7	2	2	15
Related concepts	14	16	17	1	3	50
Pre-requisites	8	5	7	0	0	20

General LM understanding under the view:

Inspection of data, illustrated in Figure IV.10, shows the “Lectures” view is the LM most likely to be understood within the four considered, with 77% of the models recognized to a sufficient level all tools considered. The view where the model was understood the most next is “Related Concepts” view, with 50% of the children having a good understanding of the model view.

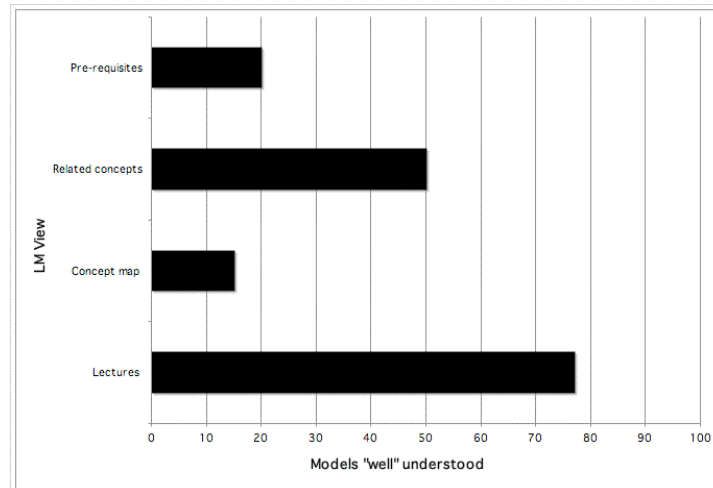


Figure IV.10: Representation of children’s understanding of learner model through a LM view.

The view of ‘concept map’ and ‘pre-requisites’, however, do not seem to facilitate children’s understanding of the LM structure and content, with respectively less than 20% and 15% of the models considered as understood to a satisfactory level.

Impact of the LM view on the understanding level using a specific visual representation tool:

There seem to be a difference in the child’s understanding of the learner model when under an LM view according to the tool used for the visualization. The two views, which scored the higher recognition of the underlying model in Table IV.4, are inspected further: the “Lectures” and “Related concepts”, illustrated respectively in Figure IV.11 and IV.12.

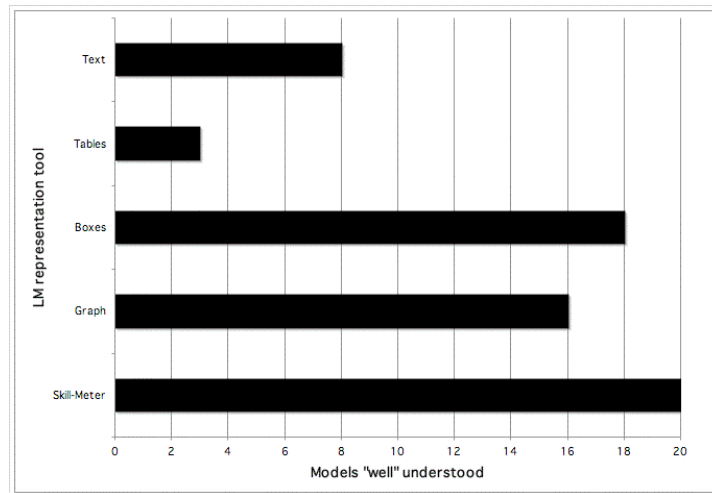


Figure IV.11: Representation of children's understanding under the "lecture" view by visual representation tool.

Inspection of the data reveals similar results to the first part of the study, also performed using this view, with "Skill-Meter" and "Boxes" the best recognized, and "Text" less recognized (40% truly recognized, and 10% with no recognition at all).

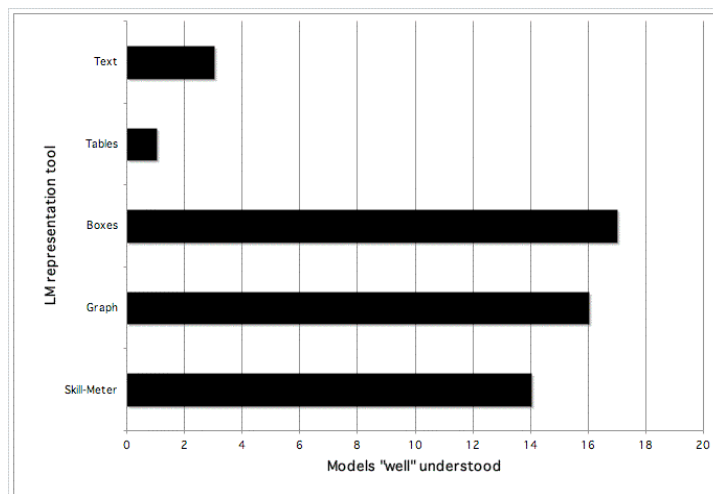


Figure IV.12: Representation of children's understanding under the "Related concepts" view by visual representation tool.

Inspection of the data reveals similarity to the observations for the "lecture view", "Skill-Meter", "Graph", and "Boxes" seem to hold the best recognized underlying model, with "Tables" and "Text" holding the lesser level of recognition (over 70% of 'week' level, less than 15% of "good" level of understanding).

CHILDREN'S PREFERENCES OF LM VIEW

Figure IV.13 pictures the representation of LM views liked the most (on the left) and least (on the right) by all children during the second part of the study.

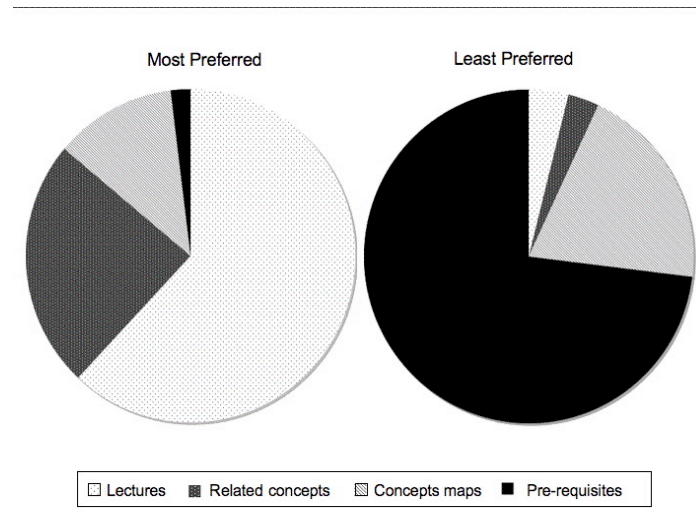


Figure IV.13: Children's preference of LM view

A two-tailed chi-square “goodness to fit” analysis performed on the rating of children's preference between visual representation tool (respectively most and least preferred) showed a significant difference in the proportion of visual representation tool chosen by children as their most preferred ($\chi^2 = 82.72$, $df = 3$, $p < 0.001$) and least preferred ($\chi^2 = 121.36$, $df = 3$, $p < 0.001$) tool. It appears that a majority of children prefer the “Lectures view”, followed closely by the “Related concepts” LM view, while they least favour the “Pre-requisites” view. Inspection of data showed that for the “concept map” and “pre-requisites” view, the ‘good’ and ‘medium’ levels of understanding all corresponded to children aged nine and more.

When applying a score from 4 (most preferred) to 1 (least preferred) to children's preferences, the following scores were obtained, presented in table IV.5:

Table IV.5: Children's rankings of LM views preferences.

Tool	(4) Love	(3)	(2)	(1) Hate
Lectures	62	30	4	4
Related concepts	24	60	13	3
Concept maps	12	5	63	20
Pre-requisites	2	3	20	73

A chi-square ‘goodness to fit’ analysis performed on the preference of the LM view showed a significant difference in children’s preference of the views, from what could have been obtained by chance alone ($\chi^2 = 121.36$, $df = 3$, $p < 0.001$). The results show a tendency for children to order their preferences in the following way, from most to least favoured: *lectures, related concepts, concept maps, pre-requisites*.

4.6 DISCUSSION

Children aged seven to eleven yrs seem more likely to favour ‘skill-meter’ and ‘boxes’ representation tools in OLM systems, and very unlikely to select Bayesian graph tools when given the choice. These results are in line with Bull & Mabbott’s (2006) observations for individual model views with “skill-meters” being the most preferred. However, whilst in Bull & Mabbott the second most preferred tool was “graphs”, usually in combination with the “skill-meters”, children in this study preferred using the “boxes” representation tool. Only 34% of the children reaching a good level of understanding of the underlying LM when using graphs, whereas 72% of them had a good understanding using the “boxes”, and 84% with “skill-meters”. The results of the understanding children gained of the learner models when presented with specific tools seem to be in line with children’s preferences. Inspection of data showed a significant difference in the level of understanding, with “skill-meters” and ‘boxes’ most likely to better help the understanding, and the “Bayesian 3D graphs” particularly hard to understand. The studies performed in this chapter ultimately aim at informing design based on the best choice of LM tools and views for OLM applications that can be used by children as young as seven years old. This last tool was therefore taken away from the choice of LM tools for the second part of the study. In Study IV.1, children subjected to ‘text’ as a LM tool achieved the lowest number of understood models. It can be observed in Study IV.2 that the change of LM structure via the views, does little to help the understanding of the models for this tool. This may suggest that the problem lies in the nature of the representation tool itself, perhaps linked to children’s ability to conceptualize textual information. It would be interesting to study the cause of this difference from other tools in further studies.

The age of the participants impacted the children’s understanding of the learner model, indicating that children seem to be more able to understand the underlying model as they develop, the extent of the understanding improvement varying according to the tool used. Two tools, the ‘skill-meter’ and ‘boxes’, exhibited a good level of understanding for all children age groups, with ‘good’ levels of understanding varying for skill-meters from 90% (seven years) to 100% (eight years and more), and for the boxes from 80% (seven yrs) to 100% (nine years and more). The eldest child population (ten and eleven yrs) appear to reach a quite good understanding for all representation tools, with the exception of the “Bayesian graph”. However, the two youngest groups (resp. seven years and

eight years) also show a significant difference of understanding reached, with two very distinct results: while the “skill-meter” and “boxes” are easily understood by children of this age, the four other categories of representation tools ranked very poorly, with more than 50% of the children having poor understanding of the underlying model.

The “best” views, both in term of children’s understanding and preference, seems to be the ‘lecture views’ and, to a lesser extent, the ‘related concepts’ view. Felder and Silverman (1995) distinguish between *sensing* and *intuitive* learners. *Sensing* learners will prefer information taken in through the senses, while *intuitive* learners will favour information arising introspectively. These two views represent the simpler lecture-oriented presentation views, with a link to the real world, usually more attractive to *sensing* learners (Mabbott & Bull, 2004).

Unlike in Mabbott & Bull’s study (2004), where no LM view was considered better than the others in term of overall usefulness, children aged seven to eleven years seem to understand better and favour views that are more sequentially organised to the “concept map” and “pre-requisites” view, which might better suit a learner who views learning as a sum of individual concepts to grasp, and will need a general understanding of the concept structure. However, it appears that while the younger children had a poor understanding of the “concept map” and “pre-requisites”, the contrary can be said for children aged nine and more. This might indicate that children, as they age, gain a more general view on learning, being able to focus on conceptual interrelationships. However, the focus of this part of the study is concerned with the relationship between understanding of the LM, the representation tool used, and the LM view the children were presented with. An analysis on the impact of age on the understanding of the LM was not performed, and no conclusion can be drawn.

Results on children’s preferences show a significant difference in their choices, ranking the views from most to least favoured. These results differ from Mabbott and Bull’s study (2004) where students favoured most the “concept map” view, and least the “related concepts” one. In this study, while students had clear individual preferences when multiple views were available, each view had a significant number of students that consider it to be most useful, and a significant number of students who considered it least useful. This does not appear to be the case in the study hereby described. Indeed, children’s preference in this study is shown in the significant difference when considering the distribution of most favoured, least favoured, and children’s ratings of the four views. The results show an association between the tools used in a LM view and the level of understanding reached for the underlying model. Children aged seven to eleven years seem more likely to understand better LM represented under a ‘lecture’ or ‘related concepts’ view, in which ‘skill-meters’, ‘graphs’ and ‘boxes’ seem more likely to help the understanding.

4.7 SUMMARY

This chapter is concerned with the design of the visual aspects of the learner model content of affective OLM tutoring systems using EPAs as an interaction medium between the child and the learner model. It aimed to answer the following research questions:

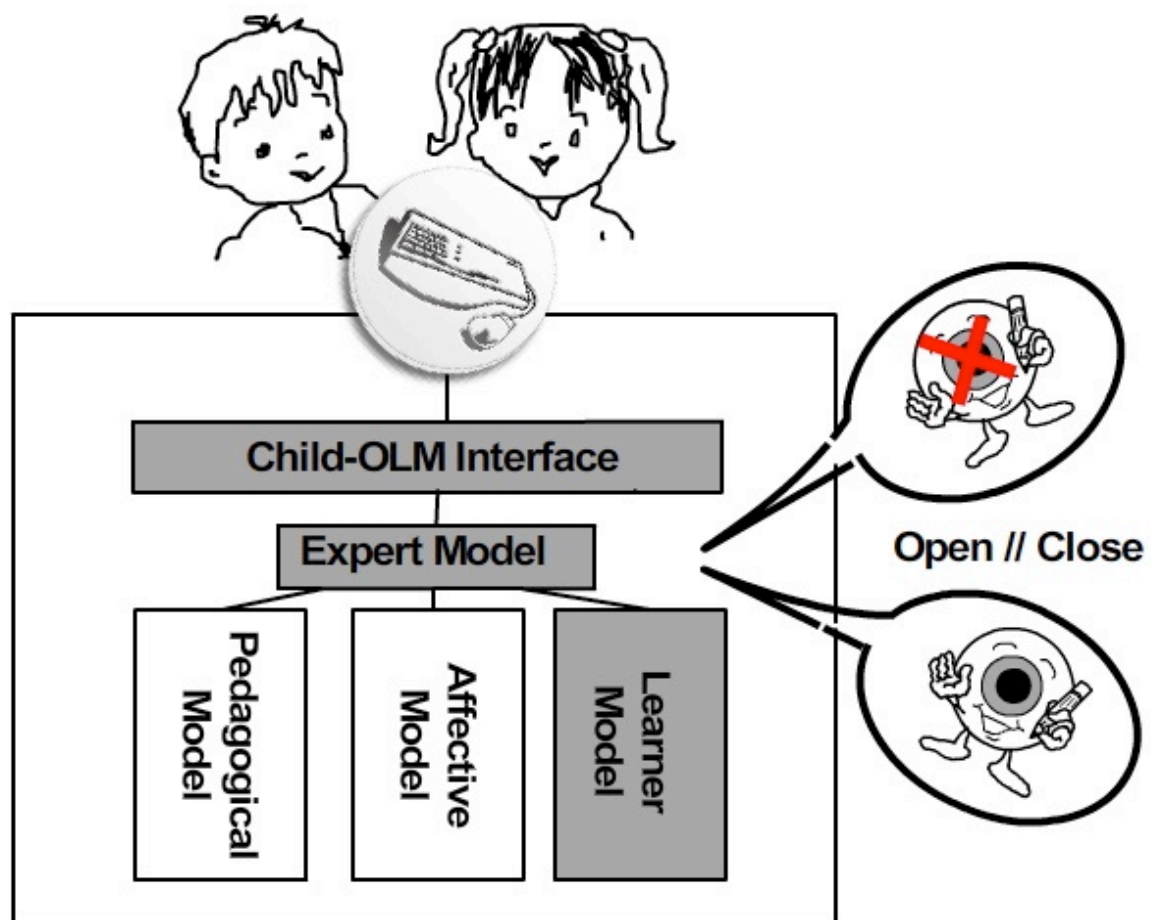
RQ3: What is an appropriate representation of a learner-model (LM) for a child to interact with?

Providing users with the most useful and easily understandable forms of LM representation is key to increasing students' willingness and ability to use OLM systems for improving their learning better (Mabbott & Bull, 2004; Kay, 1995). Multiple representations have found to be useful for students to access in OLM applications (Zapata-Riviera & Greer, 2004; Mabbott & Bull, 2004). This chapter provides some interesting results as to the understanding and preferences of use of LM representation by children aged seven to eleven years old, and were used in the following section to define the visual representations and LM views to be included in the *DividingQuest* and *Multipliotest*.

An exploratory study was performed to gain some insights on children's favoured and better understood visualization techniques to represent the learner model, producing some interesting results for French children aged seven to eleven years: The 'lectures' and 'related concepts' revealed to lead to a higher level of understanding of the model content. They also are preferred to the other tools by children this age. Under those views, there seems to be a strong preference and better understanding of the model when 'skill-meters' and 'boxes' are used as LM representation tools. The results of this study were then applied to the design of the OLM systems used in this PhD as illustrated in the design chapter (chapter 2, section 2.5.3).

Chapters 3 and 4 were concerned with the design aspects of OLM software when considering two components of the system: the affective aspects of the learner model, and the visualization of the LM itself. The results of the studies presented in said chapters were considered for the building of two OLM drill-and-practice applications, the *DividingQuest* and *Multipliotest*, which are described in details in chapter 2. Chapter 5 will now investigate a first research question concerning the use of OLM systems by children: what are the benefits of opening the LM to the user in terms of learning outcomes and child motivation, when opposed to a system where the model is hidden from its user?

Chapter 5: Open or Closed Learner Model?



5.1 INTRODUCTION

This PhD is focussed on the potential for helping children to learn better through their use of OLM software. OLMs should help the understanding of syllabus and progress, and be more motivating for children. The first step in investigating these aspects is to identify whether there are potential benefits in opening said model to child-users, and what are the issues identified with child-users from seven to eleven years of age working on OLM software.

The research issues are investigated through an experimental study of the benefits of using an Inspectable Learner model – ILM - (OLM with the model content open for inspection only, using common LM visualization tools such as boxes and skill-meters) versus ‘Closed’ Learner Model - CLM - (software where the model is hidden to the learner) for children’s learning and motivation to use pedagogical software for learning.

In the first section, the aims of the experimental study are outlined. Section 5.3 then presents the study design in terms of research hypothesis, participants, experimental design and experimental measures as well as study procedure. After a presentation of the study results (section 5.4), section 5.5 proposes a discussion section on the benefits and issues of opening the learner model to child-users.

5.2 AIMS OF THE EVALUATION

The study will investigate the benefits of using an Inspectable versus a ‘Closed’ learner model. Both versions of the software do not contain EPAs, but use traditional visual representation techniques to propose the activities and the learner model (in the ILM condition).

Aim: To Investigate the impact of the openness of the learner model on children’s learning achievements and motivation to use software. It also aims at gaining insight on whether children want the LM opened to them for inspection only, and how they would use this information.

Research questions:

- What are the effects of the openness of the learner model on learning and motivation?
- Do children want to access LM information? What use do they make of the LM information accessible to their inspection?

5.3 STUDY DESIGN

5.3.1 RESEARCH HYPOTHESES

This section presents a list of the research hypothesis analysed in this study.

Learning occurring during the session:

H_0^1 : There is no difference in prior knowledge between the children in the control or ILM conditions.

H_1 : There is a difference between the pre-test scores registered for the children in the CLM ILM conditions.

H_0^2 : There is no difference in learning gains between the closed and open learner models that students interact with.

H_2 : The children learned more during the learning session in the ILM condition than in the CLM one.

Motivation:

For the study of the level of child's motivation, the following research question was addressed using graphs and tables as illustration of data tendencies:

- Is there a difference in children's motivation to use pedagogical software between the ILM and the CLM conditions?

Interaction with the LM:

For the study of the level of children's willingness and effective use of LM information, the following research question was addressed using graphs and tables as illustration of data tendencies:

- Do children working on the CLM condition want to access more information about their LM and under which form?
- How much did children working on the ILM condition use the LM? How could the information be better represented as to facilitate data access and manipulation?

5.3.2 PARTICIPANTS

The experiment involved twenty-eight children of level “CM1” (nine to eleven years old) and their teacher from a French primary school (twelve girls and sixteen boys, average 10 years). They had practiced multiplications for eighteen months before the learning session with the software.

5.3.3 EXPERIMENTAL DESIGN

The experiment has been designed as a between-subjects comparative study with pre and post session learning tests. It includes a control and an experimental condition, differing in the access/visualization of the learner model: The control condition is a closed learner model, while the experimental condition has an open learner model where the model can be inspected and the student’s view modified. The children used either the ‘CLM’ or ‘ILM’ versions of the *Multipliotest* software, as described and illustrated in chapter 2.

5.3.4 EXPERIMENTAL MEASURES

The experimental study is designed as a between-subject comparative study across two conditions (each condition represented by a separate user-interface: one with the CLM and the other one with the ILM). It includes both quantitative and qualitative measures registered during software interaction, after completion of pre- and post-session tests/questionnaires, concerning the different research questions to be answered on learning achievements, motivation to learn, and LM component.

LEARNING GAIN

An analysis to assess whether there is a learning improvement in each condition is realized by comparing children’s pre- and post-session mathematical test results. This measure takes into account the improvement on learning on areas of the game, game activities that children may or may not have accessed during the learning session. To minimize threats to the test’s validity, child-users were encouraged to avoid guessing in general (as explained in the briefing section in 5.3.6), but especially in the pre- and post- mathematical tests, if they did not know or remember the answer to a question.

MOTIVATION

A second set of measures is taken concerning student’s motivation in using the pedagogical system for learning, and their level of engagement within learning activities during the learning session. Qualitative measures were taken on children’s motivation to use the different systems, in post-session qualitative questionnaires using Likert scales.

A questionnaire using Likert scales was used in each condition, differing in content between the children working on the CLM or ILM versions:

- Children working on the CLM version answered questions as to what type of information they would like to have access to.
- Children working on the ILM version answered questions concerning the impact the information available had on their experience using software, their understanding of the LM content, and how the information could be designed for better understanding and use.

LEARNER MODEL COMPONENT

A third set of measures is taken in the experimental condition concerning children's willingness to use, and effective use, of LM information when available to them. Children's use of the ILM was analysed in terms of number of LM inspections according to interaction patterns inspired from (Shalvrour & Bull, 2009), and the mathematical skills group defined by their teacher (Bull & McKay, 2004).

As described in details in chapter 2, the ILM condition offers the possibility for the child to inspect the content of his/her learner model, where all the levels of the *Multiplotest* are represented using skill-meters and boxes. An inspection of the learner model is registered in software when the child clicks on the inspection button (cf. Figure II.X in chapter 2), which direct him/her to the LM viewer interface. Due to the low number of concepts to acquire in this software, all information regarding the model is represented in this interface. Thus a child willing to see his/her level in, two concepts would only need to 'inspect' the learner model once.

Due to the simple and well-defined structure of the drill-and-practice application, a child can only access the learner model for inspection at specific moments of software use. These actions are recorded as software logs and can then be analyzed as a sequence of events. By taking into account the choices of interaction available to the child at each moment s/he chose to inspect the model, it is possible to extrapolate on children's intention behind model inspection, to a certain degree of confidence. This measure should therefore help give more insight on how children are willing to interact with the learner model, and what use they can find in the information the model provides.

As described in details in chapter 2, *Multipliotest* is composed of a series of activities testing children's knowledge of multiplication tables. Both the 'CLM' and 'ILM' conditions contain an interface with :

- The main screen, presenting all the activities to choose from.
- A button guiding the user to multiplication tables for revision. This button is however only accessible outside from an activity: they cannot revise a table while performing an activity on said table
- For the ILM condition, a button leading the user to the LM representation is available (cf. section 2.5.3 for details and illustration).

An activity is composed of a series of ten questions, with feedback given after the completion of each question, and at the end of the ten questions. At any point during the activity, children can see their progress by looking at the progress bar at the bottom of the page.

In both conditions, this progress bar shows the number of questions already performed and how many remain. A number of questions correctly answered appears also in the top-right corner of the activity screen. A number of questions to succeed to validate the activity (personalized for the ILM interface, and fixed for the CLM one) was also present next to the number of questions correctly answered. The ILM interface represents itself a part of the learner model in the representation of the progress bar, in the form of children's achievements is seen questions by questions. Indeed, an incorrectly answered question will be represented by a red target, crossed, and a correctly answered one will be drawn as a green shape with bold outline in the progress bar.

At last, children's actions in the ILM condition were scrutinized: the inspection of learner model content was categorized in function of children's previous actions, taken from the following list (the labels under parenthesis corresponding to the ones used in the results table):

- visualized the map of activities ("MAP")
- revised a multiplication table ("REVISION")
- abandoned an activity ("QUIT_ACT"): this is divided into activities already lost ("L"), won ("W"), or which final result is currently uncertain ("?")
- finished an activity ("FINISH_ACT"): separated according to the result of the activity ("L", or "W")

These measures were inspected with regards to the level of mathematical abilities of the participants, in order to give insights on the impact of previous knowledge on learner model use.

5.3.5 PROCEDURE

The two interfaces of *Multiplotest* version 1 were used for this study, and differ only on the learner model features represented in navigation or mathematical help, and learner model visualization/access. A complete description of the software characteristics, in particular in relation to the LM components, can be found in chapter 2. Each child worked on a unique software interface, for a single session of thirty minutes: either the “CLM”, or the “ILM” interface.

The children worked in a classroom separate from their peers, in groups of three, in the presence of a researcher. The material was conveniently placed in order for children not to see the others work, and all worked on the same experimental condition. They were therefore not aware of a difference in software interface between the children participating in the experiment until the end of the session, after completing the qualitative questionnaires.

After answering the computerized pre-test mathematical software, they were free to use software features to their liking for the rest of the allocated time. After answering the post-session computerized mathematical test, they were presented with a short questionnaire on their motivation to use software. Finally, the children working on the ILM condition were given another questionnaire on their use of the LM component.

Two days after the experimentation, the evaluators came back for a close-up session with the class. The class was divided into two groups, corresponding to the experimental conditions. They asked the children to talk about their impressions on the game, its usefulness, and improvements to be made. The children were also asked a second time to rate their willingness to use software again. This second questionnaire intended to look at children’s variability of answers when able to discuss their experience with others, across conditions.

5.3.6 PRE- AND POST- EXPERIMENT BRIEFING OF THE PARTICIPANTS.

The experiment took place in the school, with children exploring the learner models individually. When performing the experiment, the children were isolated from their peers in a specific classroom reserved for used by researchers only on this day. Three children were undertaking the experiment at the same time, on separate computers.

For the pre-task briefing, a choice was made to give the general instructions about the task to perform in front of all participants at the same time, in the form of a video projected on the classroom whiteboard. When working with child users, it is essential to make sure they understand the tasks to perform, without inferring research goals from the instructions. If not portrayed in a cautious manner, the instructions “you can do X” will be understood by most as either “you shall do X”, or “you can do X, but it would be better if you did not”, which modifies the experimental setting itself (Markopoulos et al, 2008). Had the instructions been given individually before each child, a simple smile, sign of tiredness, or contrariety on the part of the researchers at that moment could have changed the child’s entire experience with software, and therefore the study results.

A month prior to the experiment, the researcher directing the study assisted to a math lecture in the classroom, with the consent from the director of the academic centre the school belonged to, as well as the school director. She was introduced to the children, along with the experiment to be performed. Consent forms were given to the children for their parents or guardians to sign, and required on the day of the experiment to access software. The experiment was introduced to the children as an ability to practice freely their multiplication skills. The researcher participated in group activities with the children, in order for the students to familiarize themselves with her. This step was deemed essential as children aged seven to eleven can get easily distracted by elements outside from their usual learning environments (Markopoulos et al, 2008). By allowing the researcher to be viewed as “another teacher”, and not a stranger, a teacher-child relationship was formed between the children and the researcher. This facilitated children following the researcher’s instructions in a more intuitive way during the time they were alone with her for software use.

During this adaptation period, children were given a three questions questionnaire, using color-o-smileys. They were asked whether they liked maths, then multiplications, and how good they were at multiplications. This information was integrated to the teacher’s evaluation of children mathematical skills concerning multiplications: for each child, the teacher defined a set of known concepts and pedagogical expectations for software use, which helped define their learner model content. A day before the experiment, each child’s learner model was viewed by the teacher, and modified to take into account the learning that took place since the creation of the model.

The day of the experiment, children were brought to the classroom by groups of three. The order of the children using software was defined by the teacher. However, it ensured that each child within one group would work on the same experimental condition. This aimed at not letting children know about the differences in software interface across conditions until after completing the first set of questionnaires. When entering the classroom, the child was guided to one of the computers, and given the login information necessary to work on the interface. All children had some degree of familiarity with working on computers, as they practiced a few times a month learning different subjects on technology since the beginning of primary school (since the age of six).

Children were guided by the researcher to complete the pre-session mathematical test. To answer such multiplications, they were asked to think about the result and only give an answer if they were certain it was the correct one, in order to limit children choosing random answers and getting them right by luck rather than knowledge. The same process was repeated after software use, for post-task mathematical test.

Upon entering the *Multiplotest* software after completion of the pre-task test, children were given free-reign on the use of software, as long as they didn't quit the application. They were allowed to stop working on software by calling the researcher if they so wished. They then had to perform the post-task test, and could engage in drawing activities while waiting for the others to finish. Children were given the authorization to ask questions as to the usability of software, but the researcher did not answer mathematics-related questions.

At the end of the allocated time (twenty minutes), children performed the post-task mathematical test, and completed the feasibility questionnaire. Children were thanked verbally for their participation into the study, and accompanied to their classroom in order to get the next group of participants. Two days after the experiment, the class was divided into two groups, according to the experimental condition they tested. Each group was asked to complete a satisfaction questionnaire a second time. This was followed by a free discussion session on the goods and bads of software. Finally, each child was given a "certificate", as a souvenir of their participation to the project, as advertised in (Markopoulos et al, 2008).

Before leaving, the researcher thanked again all child-participants and the teacher, and gave to the class a number of personalized gold stars, bearing the logo of *Multiplotest*, as organized with the teacher beforehand. In this school, stars are earned by children upon excellent scholastic results, good behaviour, helping others etc... and they grant "special treats" for the entire class when enough are earned. This reinforced the positive experience of children in doing the experiment, and facilitated the reused of the participants for other projects ongoing, in line with the recommendations in (Markopoulos et al, 2008).

All data relevant to the experiment (questionnaires, researcher's notes, and software logs of children's exploration of software) were then anonymized and coded before being analyzed.

5.3.7 CHOICE OF STATISTICAL ANALYSIS FOR RESEARCH HYPOTHESIS

Considering the number of research hypotheses to be statistically analysed and factors of the experiment, the level of statistical significance for rejection of the null hypotheses has been defined at $\alpha = 0.01$. Two research hypothesis, H1 and H2, were analysed statistically:

- The research hypothesis H1 aims at investigating whether the distribution of students into the control and experimental conditions resulted in groups of equivalent mathematical abilities when answering the pre-session test. A chi-square “goodness to fit” statistical test was performed on the children's pre-session test scores, investigating the variance of the populations in their answers.
- The data gathered in order to investigate hypothesis H2 did not satisfy the assumption of normally distributed data, which is a requirement for any parametrical test such as ANOVA. Therefore, the analysis was performed using a non-parametric test, the Mann-Whitney two-sample rank-sum test, which tests for equality in medians.

5.4 RESULTS

LEARNING GAIN AND MOTIVATION

Learning occurring during the session:

A Chi-square ‘goodness to fit’ test failed to demonstrate a statistically reliable difference ($\chi^2=14.67$, $df=17$, $p = 0.619$) between the expected and observed frequencies in pre-test scores between the two conditions. The participants in both conditions seem to come from the same population, with a lack of statistical difference in prior knowledge across conditions. This means that inherent differences in the population would not be able to account for any differences found in the post-test, or the proportional learning gain within the session.

The results of the proportional learning gain (between the pre- and post- test scores) concerning the two conditions were compared using the Mann–Whitney two-sample rank-sum test.

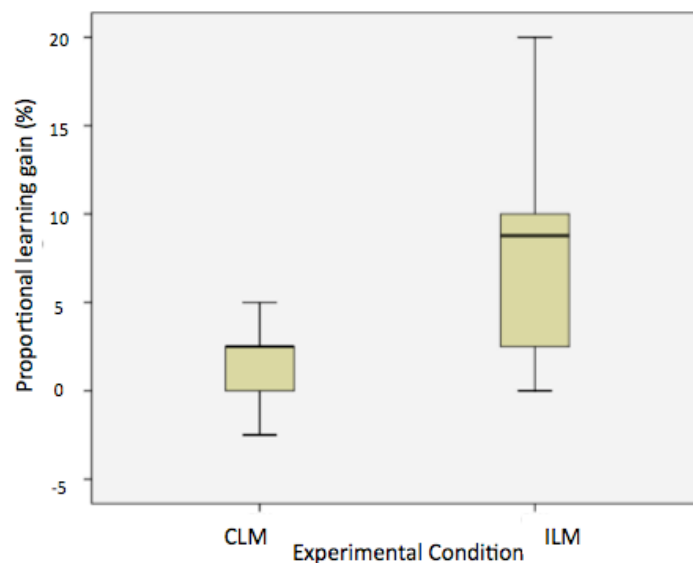


Figure V.2: Proportional Learning gain across conditions

Median for proportional learning gains in groups ILM and CLM were 0.018 and 0.077; there is a statistical difference between the learning gains observed under the two conditions (Mann–Whitney $U=31$, $n1=n2=14$, $p=0.002$ two-tailed). It can be concluded that the ILM software interface enables children to learn significantly more in terms of multiplication skills than the CLM software.

Table IV.2 illustrates the data in terms considering children's prior level in multiplications.

Table V.2: Average learning gain (%) during the session by children's level groups






	Low-skilled		Middle-skilled		High-skilled	
	Av	S.D.	Av.	S.D.	Av.	S.D.
CLM	9	1.826	5	1.301	2	0.27
ILM	20	3.509	15	3.472	5	2.35

Low-skilled children are found to be the ones benefiting more from the ILM condition (9% to 20% of learning gain between pre- and post- tests), closely followed by the mid-skilled students (5% to 15% of learning gain). While low-skilled children are also the ones benefiting most from the CLM condition, the proportional learning gain is far lesser (2% to 5%).

Motivation:

Table V.3 illustrates the answers given by the children on their motivation to use the system and level of engagement after the experiment, by means of a 5-points scale smiley-o-meter: children rated each statement, and selected a smiley that corresponded to their belief, with a scale illustrated here "does not agree" to "totally agrees".

Table V.3: Answers concerning the motivation (nb answer)

					
Statement 1: <i>Multiplotest</i> is fun to use					
CLM	3	7	1	3	0
ILM	0	1	3	2	8
Total	3	8	4	5	8
Statement 2: With the game, I learned more about multiplications.					
CLM	2	8	1	3	0
ILM	0	1	0	3	10
Total	2	9	1	6	10
Statement 3: I would like to play <i>Multiplotest</i> again					
CLM	2	9	3	0	0
ILM	0	2	3	4	5
Total	2	11	6	4	5

There appears to be a difference in children's rating tendencies concerning the level of enjoyment produced by the two different conditions: children using the ILM version seemed more satisfied with the application, experiencing more enjoyment out of the game, and visualizing more how it could be used to learn maths than the ones working on the CLM version. A majority (16/28) of children found both software interface useful to learn multiplications.

Only one child did not consider the application enjoyable in the ILM condition, and did not find any use of the application for learning multiplications. He happened to be the youngest low-skilled student of the group. When interviewed, he expressed his frustration at his inability to achieve the activities of level L2, his level in multiplications not being good enough to tackle this task.






On the contrary, the children using the CLM version of the software expressed their disappointment in not having more of a rewarding scheme and storytelling embedded in the application. They did not seem to find the application as useful as the ones working on the ILM condition. They “did not know what to choose” in the activities, and after testing a few activities, did not feel like there was enough guidance offered to them, thus leading them not to agree with statements 1 and 2. This is interesting as software content was identical in the two conditions in terms of reward and story-telling components. This might suggest that the access to the LM held enough interactivity and produced enough interest in children for them not to be focussed and bothered by the lack of ‘game-like’ features in software.

The same tendencies can be found for children’s willingness to use the application again: the majority of children using the CLM version were not too keen to use software again after the experiment, with low-skilled children being the most enthusiastic to try again. On the contrary, a majority (9/14) of the children working on the ILM interface expressed some to a great interest in using software again. No children within this condition expressed a strong dislike in software that would lead them not to use it again at all in the future.

For the next part of the questionnaire, the questions differed as a function of the experimental condition the child experimented. Table V.4 illustrates the answers of children working on the ILM condition, while table V.5 corresponds to the answers of children working on the CLM condition. During the completion of the two questionnaires, the children were not yet aware of the existence of two different versions of software.

Table V.4 shows the results of the questionnaires given to the children working on the ILM condition as to their understanding, use, and willingness to use the LM content, by means of a 5-points scale smiley-o-meter: children rated each statement, and selected a smiley that corresponded to their belief, with a scale illustrated here “does not agree” to “totally agrees”. They were presented with a screen capture of an LM as displayed in the game (defined as the “model screen” to the children), and asked to answer the statements shown in table V.4 using smiley-o-meters.

Table V.4: Questionnaire on LM use, ILM condition.

					
In the picture, the child knows its table of one.	-	-	1	2	10
In the picture, the child knows its table of seven.	10	3	1	-	-
I understand the information on the model screen.	-	3	-	1	10
The model screen helped me decide what to do next.	9	-	-	1	4
I need to look at the model screen to choose what I will do next.	9	2	-	1	2
The model screen shows what the teacher thinks I can do.	-	-	-	2	12
The model screen really shows what I know.	5	2	2	3	2

The first two questions give insight on children's understanding of the LM information: the table of one was represented as a concept fully grasped on the screen, while the table of seven was not grasped at all. While all children answered using the right polarity of answers, four children were more moderate in their answers: One child used the middle scaling point in each case, expressed his uncertainty as to the answer when interviewed. Two others used the second and fourth scaling point, stating that this may be the answer, but they weren't sure. The last child was sure of his answer for the first question, but hesitated for the second between not grasped and not fully grasped. The ratings to the third statement reflect the previous results: the children who did not systematically express a clear opinion, using scaling points other than the extremes, did not think their understanding of the LM content to be good enough. The child who hesitated for the table of seven only rated his understanding of the LM as good but not total.






The next four statements test children's willingness to use the LM and what it represents for the users: Only five children out of the fourteen considered the LM information to be useful to inform their decision as to activities to choose. The remaining nine rationalized their ratings in the following way: the information was not presented in such a way that invited learners to inspect the model and reflect on what it could mean. Three of the children expressed their need to use LM information to guide their choices and be more aware of what they knew. From the teachers' interview, these children were quite shy in class and in need of constant supervision when confronted to the learning activities. The remaining of the participants was divided between two behaviours: four of them were confident about their knowledge already, and did not feel useful or necessary to look at the information offered.

The other seven would have liked some guidance, but preferred software exploration to pausing to think about their knowledge, by using the ‘model screen’.

Children’s perception of the model content generally revolved around an assessment of their skills, performed by their teachers. Only five children felt that the learner model view represented their current level. A couple of children were undecided as they felt their understanding of the model to be limited. The remaining children had a quick glance at the model, decided the information wasn’t correct or useful, and did not return to the screen, preferring to test the activities themselves.

Table V.5 shows the results of the questionnaires given to the children working on the CLM condition by means of a 5-points scale smiley-o-meter: children rated each statement, and selected a smiley that corresponded to their belief, with a scale illustrated here “does not agree” to “totally agrees”. It investigates what type of information would children want to have access to:

Table V.5: Questionnaire on LM needs, control condition.

					
I want to see what I do not know yet.					
	2	-	-	3	9
I want to see what I do already know.					
	1	-	-	4	9
I want to see what mistakes I make					
	2	-	-	4	8
I want to show what I can do.					
	-	1	2	2	9
I want software to help me decide what to do.					
	-	-	2	2	10

Children expressed their willingness to access LM information about different elements: what they already know, the concepts they are still struggling with, and the mistakes they currently and constantly make while using software. Children from the high-skilled group, who became disinterested by software mainly due to the activities found not challenging enough, rated negatively the three first statements. When interviewed, they expressed that they did not see the advantage of representing that they knew everything perfectly. The other children working on the CLM condition, however, seemed keen on having more information on the three aspects of their current level in learning multiplications: concepts they understood, problematic areas, and misconceptions.

The second part of the questionnaire investigated children’s purpose in manipulating the LM: a majority of children wished to use the information to demonstrate their knowledge, and have some guidance on the choices they have to make, and the activities that would be better suited for them.

In the post-session interviews, which took place two days after software use therefore leaving the opportunity for children to talk about the experiment outside from the session. Children were asked to complete the questionnaire presented in table IV.3 again. There was a tendency in children who used the CLM version to change statement 3's ratings in the following way: "I would like to use *Multiplotest* again, but only the way the other group did." Indeed, children had apparently talked about their experience between the experiment and the interview, and the enthusiasm of the children using the ILM version seemed to have encouraged the others to give another try to the software, but in the ILM condition.

Finally, all children were asked a free-answer question: *How do you think the information would be best represented for you to understand in the easiest way?* For the children working on the ILM, the main criteria to be modified in the model seem to concern the interaction with the model: the information was found too static, lacking in interactivity and usual attraction. Children from both conditions suggested using a friend, a character that could guide them within the model. It was also suggested to enable access to the different activities directly through the model. Some of the children expressed their wish to "dialogue" with the model in some way to show why they thought the information is not correct on a particular topic, or ask for more details on what they are doing incorrectly.

Interaction with the LM:

Table V.6 shows children's inspection of their LM in the ILM condition according to their level of mathematical abilities, and prior action to inspection:

Table V.6: children's LM inspection according to interaction patterns and level of child's mathematical abilities

		Low-skilled					Mid-skilled				High-skilled				
		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14
MAP		1	1	0	2	0	0	0	0	1	0	0	0	0	0
REVISION		0	0	0	0	1	0	0	0	0	1	0	0	0	1
QUIT ACT		0	0	0	1	0	0	0	0	0	1	0	1	0	0
	L	0	0	0	1	0	0	0	0	0	0	0	0	0	0
	W	0	0	0	0	0	0	0	0	0	1	0	1	0	0
	?	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FINISH ACT		1	2	0	3	0	0	0	0	2	0	0	0	0	0
	L	1	1	0	2	0	0	0	0	1	0	0	0	0	0
	W	0	1	0	1	0	0	0	0	1	0	0	0	0	0
TOTAL		2	3	1	5	1	1	1	1	3	3	0	2	1	2

Children from the low-skilled group seem to inspect the LM more than the mid-skilled or high-skilled children. We can see that all children - with the exception of S11 - inspected the learner model at least once, usually from the main map of activity. Participant S11 was the top performing student in her maths class in terms of learning multiplications. She was confident in what she did or did not know about the subject. After going from an activity of level 1 to another, except from the table of one judged too easy, she was finishing the level 2 activity at the end of the session.

It is not possible to infer children's intent for this first glance at the learner model but it raises the following questions: did they visualize the information to inform their choice of activity? Or was this a part of an exploratory effect on software concerning the button leading to the model? Twelve out of the thirteen first inspections of the model occurred within the first few minutes of interaction on software, which tends to give value to the exploratory effect theory. It would be interesting to investigate intent of learner model inspection in longitudinal study, so as to separate exploration effects from the real use of the information.

Three of the children (S4, S10 and S12) chose to access the model after abandoning an activity. All inspections were a result of children knowing already what the result of their performance on the activity would be, either won or lost. Confronted to an activity already lost, S4 inspected the model before choosing an "easier" activity in the map. On the contrary, upon inspecting the model after securing a "won" activity, participants S10 and S12 chose activities relating to tables that were marked as not mastered yet for future action. One interesting fact is that all other actions performed when abandoning an activity in the ILM condition were defined as revising, before the child attempted the same activity again, or tried a different one.

While the model information was accessed a few times after the completion of an activity ("Finish-Act") by some low-skilled and mid-skilled children, this was seldom used by most of the participants. In this study, the model seemed to be more accessed when an activity was lost than won. However, due to the limited number of participants, no generalization can be done on this aspect. Participants S2 and S9, however, showed a systematic approach in using the LM information: they first looked at the LM content, and chose an activity. At the end of each activity, they systematically inspected the LM once again, in order to choose the next activity to perform. They first both choose an activity that was indicated in the LM as not completely understood.

Once S2 won this activity, he then chose a more difficult activity, also flagged as not completely understood. After loosing on this activity, he went back to inspect the model and selected an activity of lower difficulty, in level 1. S9 lost this first activity, and after inspection of the model, chose an activity deemed more accessible to her. After winning this particular activity, and inspecting the model once more, she chose to try the corresponding activity of higher level, in level 1.

5.5 DISCUSSION

The study investigated the benefits of using an ‘Inspectable’ versus a ‘Closed’ learner model. The results presented here give some evidence as to the utility of opening the learner model to child users to help children learn better, and be more motivated in learning mathematics with technology.

- *What are the effects of the openness of the learner model on learning and motivation?*

While children taking part in the experiment were drawn from the same population between conditions, they performed differently while using the software: the ones working on the ILM interface seemed to perform significantly better than the ones working with the CLM version. Low-skilled children seem to benefit more from using *Multipliotest* in terms of learning gain than the other students, especially in the ILM interface.

The use of the ILM seems to enable “free-flow” as defined by Bruce (2001), with children choosing the activities by using the information from the learner model to show their skills and competencies. Most of the children working on the version without learner model visualization appeared confused as to what their next action should be, and more children were observed choosing the activities at random, guessing questions, or jumping from one activity to another. However, such results should be investigated further, with children using the application in a longitudinal study to separate the impact of first time interaction with the game from children’s behaviour on longitudinal software use. Indeed, in this experiment, children used software for a period of thirty minutes in both conditions, one time only.

Children working on the ILM version of *Multipliotest* seemed more motivated to use the application again in the future than the ones working on the CLM version. They felt more concerned with what multiplication tables they knew, and how to “show the software” about their knowledge (current state, or newly acquired), and enjoyed playing with the game more.

- *Do children want to access LM information? What use do they make of the LM information accessible to their inspection?*

When interviewed on their willingness to access and use LM information, children answered in a way similar to Bull et al’s (2005) and Zappata-Riviera’s (2004) results on the use of ILM application with children and teenagers: children wish to be able to access representations of the concepts they understand, the ones they have not fully grasped yet, and misconceptions or a listing of errors they currently make.

The students who did not have access to the model wished to use the information to guide their choices, solve their problems, and gain control over the learning process. However, children

working on the ILM condition did not seem to use this information to really inform their choice. They expressed their doubts as to the model's accuracy and usefulness in its current state, and asked for more control over the model construction.

There seem to be a difference in children's interest in the ILM: children from the low-skilled group seem to inspect their LM more than the mid-skilled and high-skilled children. Those results seem contradictory to Bull & McKay's findings (2004) on children aged eight and nine using the Subtraction Master: in their study, the higher-ranked children seemed to be more interested in using the LM information. Here, while three out of five children inspected their model more than once in the low-skilled group, only a couple of children did so in the two other groups. Only two children out of the fourteen inspected their model intensively, referring back to their model to monitor their progress.

The children were considered performing an inspection of the LM when clicking on the inspection button and going on the LM viewer. A choice was made not to prompt the children to use this button in the pre-session briefing. Indeed, as the experiment was comparing children's behaviours when confronted to a hidden or inspectable learner model, it would have induced a bias in the experiment. Taking this in consideration, thirteen out of the fourteen children made at least one inspection. Some of those inspections could be attributed to an exploration period from children. However, only one button lead to the inspection, and the game interface was simplified as to not contain many buttons. From the notes of the researcher, who studied children's behaviours unobtrusively during the experiment, the first inspection might have been part of the exploratory process, when performed in the few first minutes. However, from the second inspection, the actions of children were more purposeful: children expressed their curiosity at inspecting the model after performing an activity to make sure that the value had changed, or inspection performed in hope to gain some insight on what might be the better activity to choose next.

The main wish children had concerning LM representation improvement concerned the interaction method with the LM and how the information was presented: children wished to be guided through LM use by software component, the usefulness of the information made clearer, and the link between the model and the activities made more explicit.

5.6 SUMMARY

This chapter is concerned with investigating the benefits of letting users access the information of the learner model. It aimed at answering the following research question:

RQ5: Does opening the content of the learner model to the child-user facilitate learning and motivation?

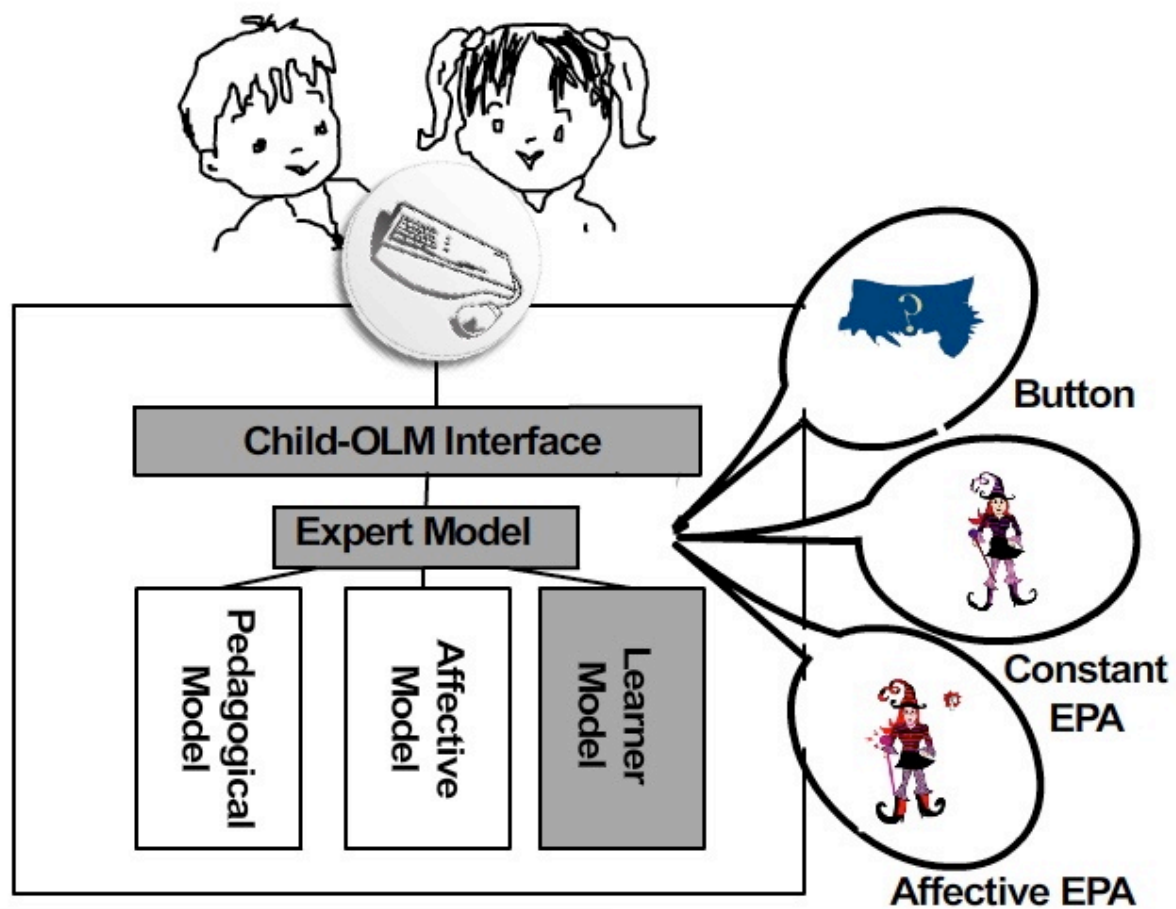
The results of the experimental study presented in this chapter show that opening the content of the learner model for inspection only to the child seems beneficial for children in terms of learning gain, motivation to use the software, and can be summarized as follow:

- Pedagogical applications which hide the learner content to the user (CLM) are less beneficial for children, in particular for low-skilled children, than a piece of software enabling them to inspect the learner content to the child-users (ILM), in terms of learning gain, and motivation to learn using technology.

Children seemed to wish to access, use, and edit the learner model information. However, the access to the LM information seemed to be used by low-skilled children only, with few children referring to the model for guidance as a systematic approach. The main concern when using the LM seemed to resolve around the means to interact with the LM and understand how grateful its use could be for learning.

Research studies have been looking at the use of Embodied Pedagogical Agents to facilitate a negotiation process in the definition of the content of the learner model between the user and the system (Zapatta-Riviera, 2003; Bull & Mabbott, 2006). In these studies, the personas, harbouring the role of Embodied Pedagogical Agent, are in charge of making children more aware of the learner model, its content, and the process by which it can be modified to incorporate student's beliefs. The visualization of the model itself is represented using traditional techniques such as skill-meters (Bull et al, 2005), or Bayesian graphs (Zapatta-Riviera, 2003). The use of pedagogical agents in the negotiation process produced promising results to improve model accuracy and children's level of reflection (Zapatta-Riviera, 2004). However, some contradictory results have been reported as to the use of said models by children, with some of them not using the models offered at all, or misusing them because of a lack of understanding of the content portrayed (Bull & McKay, 2004). In order to overcome this problem, the next chapter proposes a study investigating the potential of using EPAs portraying different emotional states to act as an interaction medium with the learner model, when embedded in the actual design of the LM representations.

Chapter 6: Affective EPAs as LM-Interaction Medium



6.1 INTRODUCTION

The experiments reported in chapter 5 revealed that opening the content of the learner model to child-users can be beneficial for learning, and children seemed keen to access more information about their current knowledge. It also showed a low level of use of the learner model information, and that some children had trouble understanding the information provided when using traditional LM representations. The use of Embodied Pedagogical Agents (EPA) in educational software has the potential to influence both students' transfer of learning (Baylor & Kim, 2005), and their self-confidence when related to problem solving (Baylor, 2005). They have also been shown to have a positive affect on students' motivation (Gulz, 2004). Using such agents to help children visualize, understand, and manipulate OLM content might increase their motivation to use the LM components for better learning gains, and be more involved in the construction of their LM representation.

This chapter aims to investigate the potential value of using an affective EPA as a main interaction medium between the system and the user to facilitate the understanding, and manipulation, of the LM by children. The research issues are investigated through an experimental study evaluating the impact of using Affective EPAs as an interaction medium with the learner model of OLM systems. It includes a control group (software without EPA, using a “traditional” representation of learner models as described in chapter 4, with only buttons as interaction with the model), and two experimental conditions (EPA with no colours and a constantly happy expression, and an EPA using both affect and colours to represent the model).

In section 6.2, the aims of the experimental study are outlined. Section 6.3 describes the experimental condition including changes in affective representations; and discusses in detail the characteristics of the OLM used, *DividingQuest*, in its three versions. Section 6.4 then presents the study design and research hypothesis tested. After a presentation of the experimental results (section 6.5), section 6.6 proposes a discussion section on the benefits on using affective EPAs to help children understand the LM information better.

6.2 AIMS OF THE EVALUATION

This study will investigate the children's ability to visualize their knowledge acquisition, and ask for help at appropriate times through the visualization of their learner model, and the interaction with the EPAs. Two sets of research issues are investigated, separated by research areas: affective EPAs and OLM issues.

Affective EPAs and Learning:

User interfaces are usually designed with traditional interface components such as buttons, slide bars, widgets, or menus. The user feedback is generally presented by written or oral instruction using such components. However, research in pedagogical agents and affective computing have shown potential in motivating the user of intelligent tutoring systems, as discussed in the literature review section (section 1.4.4). Indeed, studies have shown the successful insertion of agents in software in creating a bond between the child and the agent. The association of child's values and desired to the agent's has proved to improve motivation to use software. Furthermore, the results of chapter 5 suggest children's willingness for more interaction with software via some sort of pedagogical agent, to reduce the gap between the reality of learning mathematical concepts, and the artificiality of the software interface.

Aim: To Investigate whether there is a difference in learning achievement and motivation between a help system represented by traditional interface components (buttons, widgets) – ILM_T control condition - , an EPA bearing a single positive emotion – IML_{PA} experimental condition - , and EPAs expressing different affective states according to the user's interaction with software – AILM experimental condition. The third condition (AILM) will be represented using both colours and facial/bodily expressions in the design of the agents, while the second will have an EPA with a constant 'happy' facial expression, and no colour changes in the character's clothes.

Research questions:

- Are children more motivated to use OLM software when helped with an EPA they can associate with?
- Does the inclusion of coloured feedback and bodily/facial expressions in the EPA design help children associate more with the characters and understand better the LM representation as well as its applications?
- What are the effects of the EPAs also displaying emotions on learning and motivation?

Open-Learner Modelling:

Aim: To Investigate the impact of the use of affective and coloured EPAs, EPAs with no changes in emotional state, or buttons-based feedback, have on children's understanding and use of the software components in an inspectable LM.

Research questions:

- How much do children access the learner model to inspect it and modify its content?
- Do children perceive the LM elements displayed through the pedagogical targets and the help system? What is their impact on the inspection of their skills as portrayed by software and edition of their own representation of achievements?

6.3 STUDY CONSIDERATIONS

6.3.1 THE AFFECTIVE EXPERIMENTAL CONDITION

The participatory design process leading to the creation of the *Sorémo* self-reporting tool (cf. section 2.4.2) highlighted the possible use of coloured feedback in the design of digital characters, when combined to bodily/facial changes in expressions, to better represent affect to young users. The study presented in Appendix B also demonstrated that children aged eight to eleven recognize graphical expressions of affective states in digital characters better when the design of the characters include both colours and expressions: in general, the presence of colours (following the Traffic-Light system metaphor) in the digital representation of the affective state, reinforced the understanding of an emotional state, especially when considering emotional state of negative valence.

In the study presented in this chapter, the second experimental condition aims at manipulating EPAs's whole emotional state display according to software interaction changes. In order to test the potential of affective EPAs in the design of OLM software, a research question has first to be answered: is affect better represented by facial/bodily expressions or when the expressions are combined with a coloured feedback?

A preliminary experiment, involving forty-two children from a French primary school, aged seven to eleven years, investigated this research question. The study, presented in Appendix J, included:

- A control condition with an affective EPA designed with facial/bodily expressions only,
- An experimental condition combining facial/bodily expressions and coloured feedback following the traffic-light system metaphor to represent affect.

The feedback of the EPAs followed the A.M.E.R. model (cf. section 2.4.2), and were embedded in the OLM *Multiplotest* software, in the version of the inspectable learner model software to be described in chapter 7. The results of the study (cf. appendix J) showed that children working on the affective and coloured interface learned more and were more motivated to use software than the control group. They inspected more the LM content when looking for advice as to the next action to engage in, and were more argumentative with the characters in the feedback given.

In light of the experiment's results, a design choice was made to represent the affective experimental condition for this chapter's study by inserting both facial/bodily expressions and coloured feedback in the design of the *DividingQuest* EPA. The participatory-design process, that led to the creation of the EPA's emotional state display, was presented and discussed in section 3.3.

6.3.2 OLM AND STUDY CONSIDERATIONS

For this study, the *DividingQuest* software was used. This application, as described in details in chapter 2 (section 2.4 and 2.5), is an Inspectable Learner Model (ILM) where children can inspect their learner model information at will. The interface each child has access to is one among the following three, as directed by the child's association to one particular experimental condition: ILM_T, ILM_{PA}, and AILM. The interfaces represent the same underlying learner model, but differ in the medium provided to children to interact with the learner model content and help system:

ILM_t interface: this interface is similar to the 'ILM' condition described in chapter 5. The help system and inspection of the LM views are represented by traditional LM representation tools (skill-meters and boxes inspired pedagogical targets). The interaction between components are performed via the use of traditional user interface components, such as buttons, slide-bars, or menus.

ILM_{PA} interface: An embodied pedagogical agent (EPA), the witch 'Séléné', was added to the ILM_T interface to serve as interaction medium with the learner model content. Present in the interface throughout the game, it guides the user to the learner model display, or the help system when prompted by the child-user. This EPA harbours a constant facial display in all interactions in the form of a contented smile.

AILM: This interface is identical to the ILM_{PA} one in terms of interaction with the model. The only difference resides in the agent's representation and feedback: the agent is designed as portraying emotions, according to the A.M.E.R. model described in chapter 2 (section 2.3).

In chapter 5, the data was analysed according to the prior level of mathematical abilities of children. It aimed at investigating whether the use of ILM software could be more beneficial for children's learning than when they do not have access to the model information. In this study, while there are three different software interfaces as conditions, they all represent the same underlying learner model. In the preliminary experiment exposed in Appendix J, no significant difference of learning gains or LM use could be found between the conditions when considering mathematical abilities. As the experimental setup was similar to the study hereby presented, with an underlying learner model identical in every condition, children's prior skills in the subject of divisions will not be taken into account as a factor of this experiment's design.

The main difference between the ILM_T interface of the *DividingQuest* and the ILM interface of the *Multipliotest* lies in the measure of self-identifications: before and after software use, the children are asked to rate their own perception of their skills using smiley-o-meters. This information, when compared with the learner model content (initialized by teachers and updated according to the child's achievements on software), should help give some insight of the 'accuracy' of the learner model, following Kerly & Bull's (2006) investigations. The information is later referred to as the child's view of his/her achievements.

6.4 STUDY DESIGN

6.4.1 RESEARCH HYPOTHESIS

This section presents a list of the research hypothesis analyzed in this experiment:

Learning occurring during the session:

H^0_1 : There is no difference in prior knowledge between the three conditions.

H_1 : There is a difference between the pre-test scores registered for the three conditions (ILM_T , ILM_{PA} , or AILM).

H^0_2 : There is no difference in the proportional learning gains in the three software interfaces.

H₂: Children learn more during the learning session in the AILM condition than in the ILM_T and the ILM_{PA} ones.

Motivation:

For the study of the level of child's motivation, the following research question was addressed using graphs and tables as illustration of data tendencies:

- Is there a difference in children's motivation to use pedagogical software between the three conditions: ILM_T, ILM_{PA}, or AILM interfaces?

OPEN-LEARNER MODELLING

Interaction with the model:

H⁰₃: There is no difference in the number of times children inspected the LM between the three conditions.

H₃: Children inspected their learner model more in the AILM condition than in the ILM_T or ILM_{PA}, one.

Accuracy of the model:

H₄ is concerned with the disagreement between the system's LM view and the children's self-beliefs about their current knowledge. It computes how many disagreements were found and compare them across conditions.

H⁰₄: There is no difference in the number activities presenting LM discrepancies between the three conditions.

H₄: When considering the difference in discrepancies between the pre- and post- session models, fewer topics were found in discrepancy between the child's and system's view of the learner model in the AILM condition than in the ILM_T or ILM_{PA} one.

6.4.2 PARTICIPANTS

The experiment involved twenty-seven 'year 6' children (ten to eleven years old) and teachers from an English primary school. They all practiced divisions for six months before the learning session with the software. The children and teachers were briefed following a procedure similar to the one described in section 5.3.6.

6.4.3 EXPERIMENTAL DESIGN

The experiment was designed as a between-subjects comparative study with pre and post session learning tests. It included a control (ILM_T) and two experimental conditions (ILM_{PA} and AILM), differing in the representation of the help system:

- ILM_T: The help system of the control condition was represented by traditional user interface components, whereas both experimental conditions included an interface personas.
- ILM_{PA}: The first experimental condition's EPA presented a constantly happy expression.
- AILM: The second experimental condition included affective information in the design of the pedagogical agents, with the helping characters displaying emotions and coloured feedback according to its learner's input and achievement (cf. section 6.3.1).

6.4.4 EXPERIMENTAL MEASURES

The experimental study is designed as between-subject comparative studies across the various conditions (each condition represented by a separate user-interface). It includes both quantitative and qualitative measures registered during software interaction, or after completion of pre- and post-session tests/questionnaires, concerning the different research questions to be answered on:

- learning achievements,
- motivation to learn, and
- understanding of/interaction with the learner model components provided.

To minimize threats to the test's validity, child-users were encouraged to avoid guessing in general, but especially in the pre- and post- mathematical tests, if they did not know or remember the answer to a question.

LEARNING GAIN

An analysis to assess whether there is a learning improvement during the time spent on software in each condition is realized by comparing children's pre- and post-session mathematical test results. This measure takes into account the improvement in learning on areas of the game/ game activities that children may or may not have accessed during the learning session.

MOTIVATION

A second set of measures is taken concerning student's motivation in using the pedagogical system for learning, and their level of engagement within learning activities during the learning session. Measures were also taken on children's motivation to use the system in post-session questionnaires using Likert scales.

The measures taken with regards to the openness of the learner model are concerned with two aspects of the model:

- children's willingness to interact with the model by inspecting or modifying its content;
- the accuracy of the learner model formed (both measured quantitatively through software interaction and data logs);

LEARNER MODEL INTERACTION

Interacting with the learner model

One measure was taken to investigate the willingness of children to inspect the learner model:

- Number of times the user inspected the LM

Accuracy of the model

When concerned with the accuracy of the model, the measures concerned the results of the pre- and post-session self-identification questionnaires completed by the children, and the system's view of the student's knowledge level:

- A system's LM was defined by the teachers according to children's past performances on school tests before the learning session. A new model was created through usage of software according to children's current achievements in the different activities.
- Children were asked to complete pre- and post- session self-identification questionnaires : for each skill to be studied in software, they had to rate their current knowledge using a smiley-o-meter.

The decision concerning the establishment of the system's view of children's achievements were evaluated empirically before the study described in this chapter, and later considered as an "accurate" representation of the current child's level. In order to validate the correspondence between the system's LM adjustment to children's achievements, and teacher's views of learner model evolution, a pilot study was conducted with three teachers and twelve children (four low-skilled, four mid-skills, and four highly-skilled) from another class than the children considered for the studies in this chapter. The teachers were given the system's initial assessment of the child's level, the final model produced by the system, and a visualization of the child's interaction with the software. They were asked to check the accuracy of the model, and to annotate where some modifications would be beneficial. The models created by the system were considered accurate at 95 to 100%, with strong agreement between the different teachers on the final models (Cohen's Kappa = 0.87).

The measures calculated the agreement between the student's view and the system's view of the child's knowledge level:

- Number of topics in LM discrepancies found across all the activities between the pre- and post- test.

The measure of accuracy, introduced in (Kerly & Bull, 2006), represents how much the system's and the child's view of the learner model become identical, after using the application: it quantifies the number of topics that were found in disagreement in the pre-test, and then were agreed on in the post-test. The higher this number reveals to be, the better the model is deemed "accurate".

6.4.5 PROCEDURE

The benefits for children considered in such hypotheses concern their perception of the model, ability/willingness to use the LM, as well as the impact on learning, and motivation. A complete description of the software characteristics, in particular in relation to the OLM components, can be found in chapter 2, section 2.4.

Children were randomly assigned to one group during the class registration process in the teacher's interface according to their knowledge of divisions, in order to have children with similar mathematical abilities across conditions. They used the software for twenty minutes, in one single session for each condition, and completed a mathematical pre and post-test. They also completed a qualitative questionnaire on the use and usability of the software.

A day after the experimentation, the evaluators came back for a close-up session with the class. The class was divided into three groups, corresponding to the experimental conditions. They asked the children to talk about their impressions on the game, its usefulness, and defaults. The children were also asked a second time to rate their willingness to use software again.

6.4.6 CHOICE OF STATISTICAL ANALYSIS FOR RESEARCH HYPOTHESIS

Considering the number of research hypotheses to be statistically analysed and factors of the experiment, the level of statistical significance for rejection of the null hypotheses has been defined at $\alpha = 0.05$. The research hypothesis, H1 through H4, were analysed statistically:

- The research hypothesis H1 aims at investigating whether the distribution of students into the control and experimental conditions resulted in groups of equivalent mathematical abilities when answering the pre-session test. A chi-square “goodness to fit” statistical test was performed on the children’s pre-session test scores, investigating the variance of the populations in their answers.
- The data gathered in order to investigate the other hypotheses did not satisfy the assumption of normally distributed data, which is a requirement for any parametrical test such as ANOVA. Therefore, the analysis was performed using a non-parametric test for more than two independent samples, the Kruskal-Wallis test, which tests for equality in medians. When a significant difference was found, a post-hoc Tukey multiple comparison test was performed, in order to highlight any differences between two of the samples.

6.5 RESULTS

LEARNING GAIN

Learning occurring during the session:

A Chi-square ‘goodness to fit’ test failed to demonstrate a statistically reliable difference ($\chi^2=33.00$, $df=36$, $p = 0.612$) between the expected and observed frequencies in pre-test scores between the three conditions. The participants in both conditions seem to come from the same population, with a lack of statistical difference in prior knowledge across conditions. This means that inherent differences in the population would not be able to account for any differences found in the post-test, or the proportional learning gain within the session.






The results of the proportional learning gain concerning the three conditions were compared using the Kruskal-Wallis Test. A statistically reliable difference between the different conditions was found with $\alpha = 0.05$ ($H(2)= 7.308$, $p=0.026$) with a mean rank of 8.56 for the ILM_T interface, 15.11 for ILM_{PA}, and 18.33 for AILM. $H_{\#}^0$ can be rejected and it can be concluded that children benefited more from the AILM condition in terms of how much they learned in the session than the ones working on the ILM_T and ILM_{PA} conditions.

All the proportional learning gain scores range between -0.05 and 0.10 (Mean=0.017, Mean=0.067, Mean=0.092). However, a few 15% improvements can be seen in conditions (two children) and (one child). A post-hoc tukey multiple comparison test revealed that the ILM_T interface led to less proportional learning gains than the AILM interface ($q= 4.372$, $p < 0.05$). however no difference could be found in learning gains between the ILM_{PA} and AILM ($q=2.301$, $p=0.12$), or the ILM_T and ILM_{PA} interfaces ($q=2.107$, $p=0.08$).

Motivation:

Table VI.3 illustrates the answers given by the children across all conditions on their motivation to use the system and level of engagement after the experiment, by means of a 5-points smiley-o-meter: children rated each statement, and selected a smiley that corresponded to their belief, with a scale illustrated here from “does not agree” to “totally agrees”.

Table VI.3: Answers concerning the motivation (nb answer)

					
<i>Statement 1: DividingQuest is fun to use</i>					
ILM _T	0	5	2	1	0
ILM _{PA}	0	2	1	3	2
AILM	0	1	0	3	5
Total	0	8	3	7	7
<i>Statement 2: With the game, I learned more about divisions.</i>					
ILM _T	0	4	3	1	0
ILM _{PA}	0	1	2	3	2
AILM	0	0	1	3	5
Total	0	5	6	7	7
<i>Statement 3: I would like to play DividingQuest again.</i>					
ILM _T	1	4	2	1	0
ILM _{PA}	0	2	1	3	2
AILM	0	2	0	2	5
Total	1	7	3	6	7


There seems to be a strong difference in children’s rating tendencies concerning the level of enjoyment produced by the three different conditions: children using the ILM_T version seemed less satisfied with the application, taking less enjoyment out of the game, and visualizing less how it could be used to learn maths than the ones working on the versions with EPAs. In the concluding free-expression session, they expressed their regret at the lack of interactivity within the LM visualization interfaces of the software, compared to current pedagogical games, some of them describing the game as boring because of it. Most of the children testing this condition did not read the instructions given, and therefore felt “lost” when deciding which action to do next.

On the contrary, the children using the ILM_{PA} and AILM versions seemed to enjoy playing the game. When reviewing their learning session in the post-test interviews, they expressed their willingness to follow what the EPA was saying or advising. Fifteen out of the eighteen children described the character as a friend or a mentor; someone who knew more about divisions than they did, and whose purpose was to help them learn better. Some of them did not agree with the helper’s assessment of their knowledge level, but were willing to follow the advice given nevertheless. The same tendencies can be found for children’s willingness to use the application again. Again, the post-

session interviews taking place a day after the experiment, some of the children working on the ILM_T interface, who did not work with EPAs embedded in the interface, expressed their willingness to play again, but only if it included the “friend” the others had on the game.

Table VI.4 illustrates the answers given by the children who worked on the ILM_{PA} and AILM interfaces, on their perception of the EPA and its usefulness.

Table VI.4: Answers concerning the EPAs (nb answer)

					
<i>Statement 4: The witch helped me learn better.</i>					
ILM _{PA}	1	0	2	4	1
AILM	0	0	1	3	4
Total	1	0	3	7	5
<i>Statement 5: The witch's emotions change when I win or loose.</i>					
ILM _{PA}	8	0	1	0	0
AILM	0	0	1	1	7
Total	8	0	2	1	7

In the EPA versions, the character was perceived as quite helpful on average, the affective character rating a higher degree of usefulness than the ILM_{PA}. The change in affective expressions in the last condition seemed also quite well perceived: in addition to statement 5, children were asked to cite a few emotions they had observed during software use in the post-session interviews. One child from the ILM_{PA} condition rated the character as changing emotions, but could not tell the evaluator which different emotions were portrayed in the interview.

OPEN-LEARNER MODELLING

Interaction with the model:

Number of LM inspection

A Kruskal-Wallis analysis of the number of LM inspections observed revealed a statistically reliable difference between for the different conditions ($H(2) = 13.964$, $p < 0.001$) with a mean rank of 6.11 for the ILM_T interface, 16.85 for ILM_{PA}, and 19.31 for AILM. H^0_3 is rejected and it can be concluded that children statistically inspected their learner model more in the AILM condition than in the ILM_T or ILM_{PA} one. A post-hoc Tukey multiple comparison test revealed that the button-like condition results in children inspecting their LM less than the AILM interface ($q = 5.028$, $p < 0.05$). However, no statistically reliable difference in the number of LM inspection could be found between

the ILM_{PA} and AILM interfaces ($q=2.150$, $p=0.073$), or between the ILM_T and ILM_{PA} conditions ($q=2.878$, $p=0.10$).

Accuracy of the model:

Percentage of topics where an agreement was reached

A Kruskal-Wallis analysis of the number of topics where an agreement between the system's and the user's view of the LM was reached between the pre- and post- revealed a statistically reliable difference between for the different conditions ($H(2)= 15.025$, $p=0.001$) with a mean rank of 9.00 for the ILM_T interface, 12.05 for ILM_{PA}, and 22.06 for AILM. H^0_5 is rejected and it can be concluded that when considering the difference in discrepancies between the pre- and post- session models, statistically fewer topics were found in discrepancy between the child's and system's view of the learner model in the AILM condition than in the ILM_T or ILM_{PA} one.

A post-hoc Tukey multiple comparison test revealed that the button-like condition seems to lead to a statistically lower reduction in number of LM discrepancies between the pre- and post-test than the AILM interface ($q=0.508$, $p<0.001$). Similarly, the ILM_{PA} seems to lead to a statistically lower reduction in number of LM discrepancies between the pre- and post-test than the AILM interface ($q=0.411$, $p<0.001$). However, no statistically reliable difference between the pre- and post-test could be found between the interfaces ILM_T and ILM_{PA} ($q=0.097$, $p=0.437$) for this measure.

6.5 DISCUSSION

The study investigated the benefits of using an EPA with a changing emotional state (AILM), versus an EPA harbouring a constant happy expression (ILM_{PA}), versus an interface without EPA to serve as an interaction medium with the learner model (ILM_T). The results presented here give some evidence as to the utility of using an affective EPA to interact with the child about their learner model to help children learn better, and be more motivated in learning with educational application.

- *What are the effects of the EPAs also displaying emotions on learning achievements? Are children more motivated to use drill-and-practice software when helped with an EPA they can associate with?*

A statistically significant difference could be found in the learning gains across all conditions, with children from the AILM condition apparently learning more from software use than the children

in the other two conditions. However, all proportional gains are between -5% and 10%, and one session might not be enough to represent accurately a difference in learning gain tendencies. No significant difference could be found between the two interfaces with EPA, or between the ILM_T and ILM_{PA} interfaces. This could be explained by the restricted time the children spent on the software. Indeed, the pre- and post- mathematical tests evaluated children's knowledge on all the skills tested in the software, but children performed an average of 4 activities during the session out of the 15. For this reason, an investigation of learning gains when children had the time to explore all activities and test their knowledge against all skills targeted reveal whether children might indeed benefit statistically more from an AILM in terms of learning gains.

- *Do children perceive the LM elements displayed through the pedagogical targets and the help system? Do children associate with the characters? Do children perceive and appreciate the emotional aspects of the characters?*

Results from the questionnaires show that children working on the version of *DividingQuest* with EPA seemed to have enjoyed using the software, and associated at some level with the characters. In the AILM version (with an EPA changing emotions), the change in emotions was perceived by most children, and helped in keeping them motivated during the game. The EPA took the role of a mentor or a companion during their learning experience using the software, and the affective aspect of the character made the guidance given by the characters seem more believable for the children involved in the study. In the button-like interface, however, children did not appreciate working on the software as much, with few of them willing to work on this interface of the software in the future. They listened less to the oral feedback given by the software (identical to the one given in the EPA interfaces), and appeared to choose more activities at random, not taking into account the information about their state of knowledge accessible through the learner model. Children's identification with the animated character, acting as some kind of "play-prop", seemed to help them in mentally representing what they know, and to act accordingly when choosing their next action. Consequently, "free-flow" (Bruce, 2001) seems more present in the OLM application representing the interaction between the model and the learner by an EPA, as it facilitates the creation of play-props and play-pretend behaviours in children, increasing their level of motivation to use the pedagogical game.

- *How much do children access the learner model to inspect it and modify its content?*

The results concerning the number of learner model inspections made by children during the learning session show a statistically reliable difference between the three conditions: children working on the ILM_T interface inspected their LM statistically less than the other ones working on the ILM_{PA}, and AILM interfaces. Children seemed to complete the pre-session self-identifications with the same level of accuracy when compared to the system's view of the model in the three conditions. When

considering the percentage of topics in disagreement in the pre-test and where an agreement was reached between system's and child's view in the post-test, children working on the AILM seemed to disagree on their level of competence on fewer topics than the ones working on the ILM_{PA} and ILM_T interface. The reduction in discrepancies did not always correspond to a child changing his/her vote to one similar to the system's view. Indeed, some of the changes were orchestrated the following way: children identified a discrepancy on one particular topic; they attempted the corresponding activity to show their view on the level; from the results of the activity, they changed the content of their self-identification (usually when they lost), or the system's view was updated (the child showed a better understanding of the concept than what is portrayed by the software, the level increases). This suggests that a higher level of interaction between the model content and the user might be welcome by children this age, and even have the potential to help create more accurate OLM systems by increasing the user's participation in the model building.

From these observations, it can be suggested that the AILM interface may have led children to reflect more on what the learner model represented, how it was linked to the software activities, and how it could be used to help them plan their learning session.

6.6 SUMMARY

This chapter is concerned with investigating the benefits of using affective EPAs embedded with the traffic-light and smiley-face metaphors of learning, to serve as an interaction medium between the user and the learner model content, in mathematical drill-and-practice OLM software used by French and English children aged seven to eleven years old. It aimed at answering the following research questions:

RQ6: How can the use of pedagogical agents or the inclusion of affect in the design of OLM tutoring systems help children interact with their LM, facilitate learning and motivation?

RQ7: How can children interact with their learner-model?

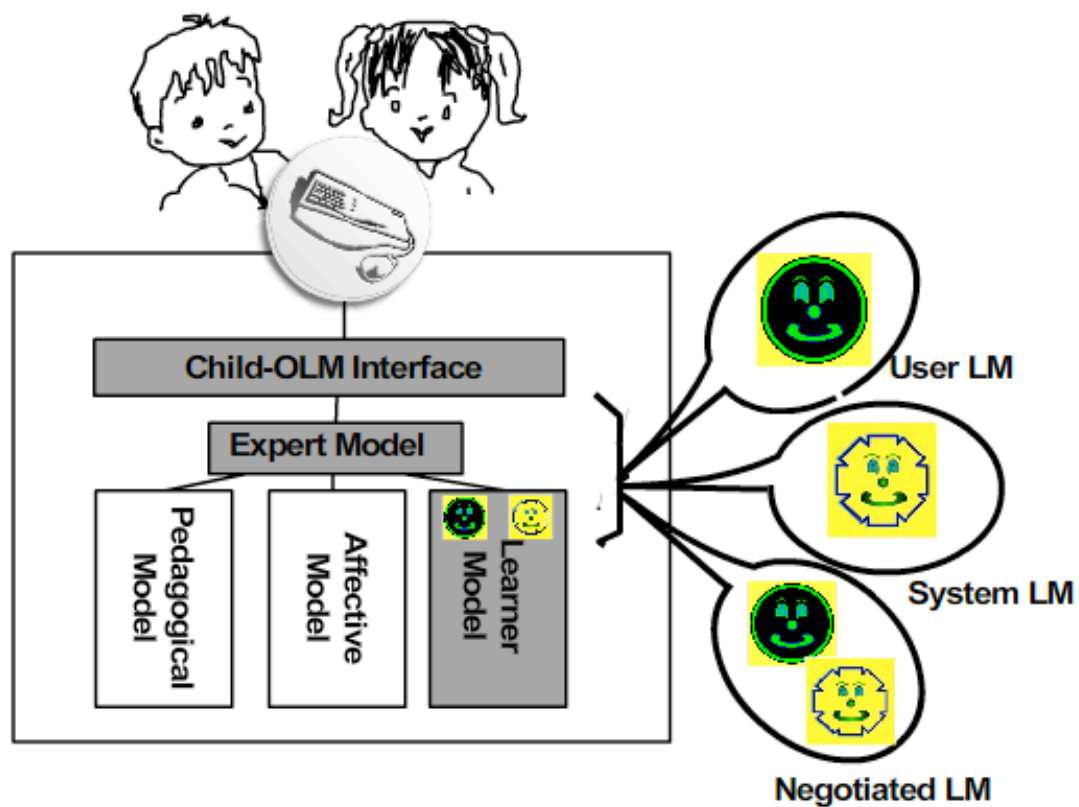
The experimental study presented in this chapter investigate children's learning gains, motivation to use pedagogical software, and understanding/use of the learner model content offered to plan their learning session and be more aware of what they know/could achieve by using the pedagogical activities proposed. The results can be summarized in the following way for all these aspects:

- Using an EPA with dynamic affective state according to the users software interaction to interact with the content of the learner model is more beneficial than using a mode “traditional” EPA (which expression does not vary from a plastered happy smile), or an interface where the interaction is only led by verbal feedback and the use of traditional UI components in terms of children's motivation to use software, the number of times the model was inspected, and the definition of a more “accurate” learner model.

As discussed in chapter 4, the learner model of intelligent tutoring systems can be opened to represent various pedagogical content, under different visual and conceptual representations. The results of the experimental studies presented in this chapter show the potential benefits of embedding affect in the design of EPAs for use in ILM systems: an EPA portraying different affective states during the elaboration of mathematical tutoring systems for children with an inspectable learner model, and serving as an interaction medium between the child-user and the learner model, can lead to better learning and children being more motivated to lean with technology.

In the next chapter, a final experiment will be presented investigating the ways in which children are willing and able to use the information presented to them via affective EPAs to gain some understanding of their knowledge level, and guide their interaction with pedagogical applications by challenging the system's view of their knowledge level state.

Chapter 7: Inspectable, User-Editable, or Negotiated Learner-Model?



7.1 INTRODUCTION

Research studies in OLM have shown that exposing the system's learner model to learners and/or teachers helps promote reflection, interactive diagnosis, collaborative creation and inspection of learner models, and provides the teacher with a powerful assessment tool (Zapata-Riviera, 2003; Bull & Kay, 2010). The experiments in chapter 5 and 6 showed that using an OLM, especially with affective an EPA serving as interaction medium, is revealed beneficial for learning is possible to increase the level of users' motivation to use the content of the learner model for better learning. Children have shown interest in inspecting the model for guidance, and "telling the software" what they thought their level was, using a representation of LM content representing their own judgement. However, some issues remain as to the model access control, and how child-users actually use the model. How much do children want to interact with the model? Can they make decisions from seeing the LM by themselves, or would they need some guidance? Do they want to take part in the construction of the LM, and to what extent?

In most 'traditional' OLM systems, the LM is represented by one view, later referred to as system's view. This corresponds to the underlying model content, and acts as the expert representation of the user's achievements and skills. Information in such application can be seen without any possibility of content modification, called Inspectable Learner Models (ILM). Others, defined as Negotiated Learner Models, enable a negotiation to take place between the user's point of view and the expert system to build the model step by step. For each concept integrated in the LM, the child enters a negotiation process with the application in order to fix the value of the corresponding LM item. This negotiation can be from directly modifying the content (User-Editable Learner Models), to processes of persuasion giving proofs, as described in the literature review, section 1.4.4. However, allowing children to modify the content of the expert model might not be beneficial for all children involved, and children will still have to compare the information given to a mental representation of what their knowledge is. This implies a huge cognitive involvement on their part (Dimitrova, 2003; Bull & McKay, 2004), which can be from difficult to impossible to children at this stage of development.

In this chapter, a new approach to OLM system use is evaluated: The learner model is the system's view of the child's achievements, and can be compared with children's own perception of their current skills. Children are able to visualize the discrepancies between their own beliefs and the system's without relying only on their capacity of mental representation. The extent to which the learner model can be opened for children to understand its components, and use them consciously when learning with computerized environment, is investigated.

A version of the *Multipliotest* software, described in chapter 2; was used with user-interfaces including different OLM aspects, and the corresponding level of control children have over the model content:

- u-ELM : An *User-Editable Learner Model* with the child's view of the learner model open for children to manipulate without directions, the system's view staying hidden from the learner at all times: control by user of their own view, no control over the system's view;
- AILM: An *Inspectable Learner Model* where the interaction with the learner model is guided by visualizing the system's view. Only the system's view is represented and can be inspected, but not edited: control by system;
- NLM: A *Negotiated learning model* comparing the two views to achieve some agreement in the model definition. Children can see each item of the LM with a representation of the system's and the child's beliefs about the knowledge of this particular item. This child can "negotiate" with software to try and reach an agreement between the two representations: mixed control;

In section 7.2, the aims of the experimental study are outlined, followed by a short description of the experimental conditions manipulated in software. Section 7.3 then presents the research hypotheses tested and study design. After a presentation of the experimental results (section 7.4), section 7.5 proposes a discussion section on the use children make of various OLM systems and its impact on learning and motivation.

7.2 AIMS OF THE EVALUATION

7.2.1 RESEARCH AIMS

This study will investigate the impact different learner model ‘openness’ has on children’s learning processes, the ability for children to interact with the model and have a good understanding of their current knowledge, and children’s reports of emotional states while using the software.

Aims:

1. To investigate whether there is a difference in learning achievement, between an OLM system presenting only the user’s view with edition rights, another presenting the system’s view available for inspection only, or a third with a negotiation process between the user’s and the system’s view of the learner model.
2. To investigate the impact the different views of the learner model offer to children, and their understanding of and use of the content learner model to plan their learner session for better learning.
3. To investigate the impact the use of different OLM representations can have on children’s willingness to report their emotional state, and the emotional content of such self-reports.

Research questions:

Learning gain:

- What are the effects of the OLM representation used on children’s learning achievements?

Open-Learner Models:

- Is there a difference in children’s willingness to use LM components, and their actual interaction with the LM content, according to the type of OLM application they are presented with?
- Is there a difference in children’s agreement with the system’s view, according to the OLM representation used?
- Is there an age difference in children’s use of the OLM components to access/modify the content of the model, and reflect on their state of knowledge?

In chapter 5, children's prior skills in math abilities were investigated as a factor to the child's use of ILM systems. In this experiment, each age group contains children of low, middle, and higher abilities. This experiment could have been designed using skills as a factor with 6 levels. However one can argue that children from the CE2 class will be less skilled than the CM1 ones as they learned the concept one year less. In this experiment, age was favoured to math skills as a factor of design. As the experiment is investigating children's ability to mentally represent their knowledge and make comparisons with others' views, consideration of age in line with the theories of age development in mental models construction seemed more appropriate.

Emotional aspects :

- Is there a difference in the child's willingness to report their emotional states during the interaction with the three software interfaces?
- Is there a difference in the emotional states reported during the interaction with the three software interfaces?
- Is there a developmental difference in children's use of the self-reporting tool, *Sorémo*, to express their emotional state?

7.2.2 OLM CONSIDERATIONS

A version of the *Multipliotest* software, described in chapter 2; was used with user-interfaces including different OLM aspects, and the corresponding level of control children have over the model content:

- u-ELM : An *User-Editable Learner Model* with the child's view of the learner model open for children to manipulate without directions, the system's view staying hidden from the learner at all times: control by user of their own view, no control over the system's view;
- AILM: An *Inspectable Learner Model* where the interaction with the learner model is guided by visualizing the system's view. Only the system's view is represented and can be inspected, but not edited: control by system;
- NLM: A *Negotiated learning model* comparing the two views to achieve some agreement in the model definition. Children can see each item of the LM with a representation of the system's and the child's beliefs about the knowledge of this particular item. This child can "negotiate" with software to try and reach an agreement between the two representations: mixed control;

7.3 STUDY DESIGN

7.3.1 RESEARCH HYPOTHESIS

This section presents a list of the research hypothesis analyzed in this experiment.

Learning occurring during the session:

H_1^0 : There is no difference in prior knowledge between the three conditions.

H_1 : There is a difference between the pre-test scores registered for all three conditions (u-ELM, AILM or NLM software interface).

H_2^0 : There is no difference in the learning gains between the three conditions.

H_2 : Children learned more during the learning session in the NLM condition than in the u-ELM or AILM one.

Interaction with the model:

H_3^0 : There is no difference in the number of times children inspected the LM between the three conditions.

H_3 : Children inspected their learner model in the NLM condition more than in the u-ELM or AILM one.

H_4^0 : There is no difference in the number of times children inspected the LM between the two age groups.

H_4 : There is a statistically reliable difference between the expected and observed frequencies of the number of LM inspections between the younger and the elder children age groups.

H_5^0 : There is no difference in the number of times children modified the content of their LM self-identifications between the three conditions.

H_5 : Children modified the content of their self-identifications LM more in the NLM condition than in the u-ELM or AILM one.

H_6^0 : There is no difference in the number of times children modified the LM between the two age groups.

H_6 : There is a statistically reliable difference between the expected and observed frequencies of the number of LM modifications between the younger and the elder children age groups.

Emotional Input

H⁰₇: There is no difference in the average number of self-reports registered for a child working on the NLM, the AILM, or the u-ELM condition.

H₇: The average number of self-reports registered for a child working on the NLM condition is statistically higher than for children working on the AILM or u-ELM conditions.

For the study of the emotional aspects of OLM use, two research questions are also investigated, using graphs and tables as illustration of data tendencies:

- Is there a difference in the number of self-reports given by a child from the eldest or the youngest group? Is there a difference in the repartition of the representation of an emotional state (single unique, succession of unique, or complex view of emotions) in self-reports given by a child from the eldest or the youngest group?
- Is there a difference in the repartition of the different emotions scaled in self-reports given by a child working on the NLM, AILM, or u-ELM condition, in terms of degree of pleasantness and arousal?

7.3.2 PARTICIPANTS AND EXPERIMENTAL DESIGN

The experiment included 90 children from four different classes and their respective teachers from a French primary school: two classes of level ‘CE2’ (forty-five children in total aged seven to nine years av. 8.5 years), and two classes of level ‘CM1’ (forty-five children in total aged nine to eleven years av. 10.2 years). The children were practising the learning of multiplications respectively for six and eighteen months before the first learning session on software.

The experiment has been designed as a between-subjects longitudinal comparative study with pre and post session learning tests, as illustrated in table VII.3:

Table VII.3: Grouping of study participants by condition and age.

Condition	CE2	CM1
u-ELM	15	15
AILM	15	15
NLM	15	15

The study includes a control and two experimental conditions, differing in the representation of, and access to the user’s learner model, as illustrated in Figure VII.1.

All conditions use pedagogical agents portraying emotions as an interaction medium with the learner model, the display of affective information within the EPA varying according to the learner's interaction and achievements.



Figure VII.1: Experiment design, the 3 OLM user-interfaces

The control group, AILM, is represented by an inspectable learner model software, with children visualizing the **system's** view of their LM. The first experimental condition, u-ELM, included a **user's** view of his/her LM only, without any access to the underlying system's view of the LM. Finally, the second experimental condition, NLM, comprised a **Negotiated Learner Model**, offering both the user's and the system's views of the LM to children.

In the example of Figure VII.1, the boxes represent children's knowledge of the one times table. In the u-ELM interface, the child rated his/her knowledge of the one times table as "very good". In the AILM interface, the system's measurement of this child's knowledge of the one times table is "very bad". Therefore, in the NLM interface, the box would illustrate the two points of view with the system's EPA ("*Multiplio*") portrayed as "very bad", and the child's EPA ("*Moiije*"), as "very good".

7.3.3 EXPERIMENTAL MEASURES

The experimental studies presented in this chapter is designed as between-subject longitudinal comparative studies across the various conditions (each condition represented by a separate user-interface), with each child using the same condition three times.

Learning gains and Open-Learner Models

The measures taken with regards to the openness of the learner model are concerned with It includes both quantitative and qualitative measures recorded during software interaction, or after completion of pre- and post-session tests/questionnaires. These measures concern the different research questions to be answered on learning achievements, and understanding of/interaction with the learner-model components provided.

In addition to the quantitative measures of learner model inspections and modifications described in chapter 6, two sets of measures were analysed: children's behaviour towards the choice of an activity when related to LM inspection; and, for the NLM condition only, children's participation in the jointly-constructed model.

Learning strategies

In order to analyse how, and when, children are willing to inspect their learner model, children's behaviour was divided into five categories from observations during the study and traces of software interaction, as illustrated in Table VII.4. This process was inspired by the one followed in studies III.1 (detailed in section 3.3.3) and V.1 (cf. section S.3.4) : from the elements available in the literature, a set of metrics was framed from the data provided by a pilot study. The data of this exploratory study was then coded according to the scheme defined, and analysed.

Table VII.4: From LM inspection to action

C1	<i>"one order once and for all"</i> : the child takes the activities always in the same order, for example from the table of 1 to the table of 10, then activities of higher difficulty. s/he goes to the next activity only when the current one is considered fully grasped.
C2	<i>"revise first, learn later"</i> : the child begins by the activities s/he knows best, before attempting the other ones, with an increasing level of difficulty
C3	<i>"the harder, the better"</i> : children try the most difficult activities first. When they fail, they inspect the model and go for the next activity in decreasing level of difficulty.
C4	<i>"proving the system wrong"</i> : the child systematically inspects the model for discrepancies between the system's view and their own evaluation of knowledge, and does the activity to demonstrate their opinion. After the activity is performed, they inspect the LM again to check the changes.
C5	<i>"other"</i> : Other behaviour, including what appears to be "random" choosing of an activity

Children's negotiation behaviour

A final measure was taken when considering the OLM research aspects: children's participation on the jointly-construction of the model in the NLM condition. When confronted with a discrepancy in the system's and the child's view, children responded in one of the following ways, illustrated in Table VII.5:

Table VII.5: Children's response to discrepancies

R1	Agrees with the system, no negotiation
R2	Disagrees with the system, no negotiation
R3	Negotiates with the system's view.
R4	Disagrees, and does the activity to "prove the system wrong"

This coding scheme was inspired by Kerly & Bull's (2005) results on NLM and created in the same manner.

Emotional Input

The measures taken with regards to the openness of the learner model are concerned with three aspects of the model: children's willingness to interact with the model by inspecting or modifying its content; the accuracy of the learner-model formed, how much children reflect on the LM and its link to learning (both measured quantitatively through software interaction and data logs); and an emotional component of the learner model: children's evolution of emotional state (from a pre- and post-session *Sorémo* report, as well as self-reports during software interaction).

People's emotional state changes with time according to the emotional stimulus they encounter (as described in the literature review, section 1.3). Children's emotions should therefore also evolve during software use. Investigating the events that led to the change in emotional state and how useful *Sorémo* is as a tool for children to express such change is of interest here.

Children's evolution of emotional state was measured from the results of a pre- and post-session *Sorémo* report, as well as self-reports during software interaction. The results were measured in terms of number of self-report given by the children, as well as the emotional content of the reports, and the way the reports were formed:

- Number of self-reports per learning session and in total
- Number of times each emotion was reported as felt during software interaction
- For each self-report and pre-/post-session reports, an emotional indicator was calculated in terms of "level of pleasantness" and "level of arousal" as described in details in section 2.3.3.

7.3.4 PROCEDURE

Children were assigned to one of the three conditions during the class registration process in the teacher's interface, in order to have children with equal mathematical abilities in each age group. Each child used the respective software interface for three sessions of thirty minutes, one month apart between sessions. They performed individually, on a computer laptop, in a classroom separate from the other classmates, with two other children performing the same activity in the room, under the same condition.

They were asked to complete several pre/post session tests: a mathematical test checking their knowledge of multiplications, learner-model self-identifications, and a *Sorémo* report on their current emotional state. Finally, a week after the last session, they completed a questionnaire on the use and usability of the OLM application tested.

Traces of all interactions between the child and the educational application were kept. They were used in a quantitative analysis to help the investigation of the link to the various components of the UM, and their effect on children's interaction with the model itself (for example, presence of reflection), their development of cognition skills, emotional state evolution, and impact on learning and motivation to work on mathematical software.

In addition to the pre/post emotional input given by children, users can choose to fill a *Sorémo* self-report questionnaire at any moment during the session by clicking on the "*Sorémo* Witch" button, embedded in all OLM software interface conditions.

The usability studies presented in Appendix G and H on the *Sorémo* method showed a practice and exploratory effect, which can be reduced after a couple of self-report completion. For this reason, a week prior to the experiment, all children participating in the study performed a pilot study using the *Sorémo* self-report method to express their emotional state at different moments of learning during classroom pedagogical activities. They had a paper version of the *Sorémo* interface, and had to follow the same rating procedure, when finishing an activity that produced a strong emotional response from them. They were prompted to ask as many questions as possible on how to complete the instrument's report, as well as the representation of the witches' emotional states, and to describe what they were doing when they felt this influx of emotions. This should help reduce the exploratory time needed on *Sorémo* to complete the first report during learning session, and help children be more familiar with the emotions included in *Sorémo* prior to the experiment.

7.3.5 CHOICE OF STATISTICAL ANALYSIS FOR RESEARCH HYPOTHESIS

Considering the number of research hypotheses to be statistically analysed and factors of the experiment, the level of statistical significance for rejection of the null hypotheses has been defined at $\alpha = 0.001$. The research hypothesis, H1 through H7, were analysed statistically:

- The research hypothesis H1 aims at investigating whether the distribution of students into the control and experimental conditions resulted in groups of equivalent mathematical abilities when answering the pre-session test. A chi-square “goodness to fit” statistical test was performed on the children’s pre-session test scores, investigating the variance of the populations in their answers.
- The research Hypothesis H4 (resp. H6) aims at investigating whether the distribution of inspections (resp. modification) of LM between the younger and elder age group differ from what could be expected by chance alone. A chi-square “goodness to fit” statistical test was performed on the children’s number of LM inspection (resp. modifications), investigating the variance of the populations in their answers.
- The data gathered in order to investigate the other hypotheses did not satisfy the assumption of normally distributed data, which is a requirement for any parametrical test such as ANOVA. Therefore, the analysis was performed using a non-parametric test for more than two independent samples, the Kruskal-Wallis test, which tests for equality in medians. When a significant difference was found, a post-hoc Tukey multiple comparison test was performed, in order to highlight any differences between two of the samples. This analysis was performed on the data separating the results for each of the three learning sessions, in order to highlight any difference in tendencies between the first software use, and further learning sessions.

7.4 RESULTS

LEARNING GAIN

Learning occurring during the session:

For the first learning session, a Chi-square ‘goodness to fit’ test failed to demonstrate a statistically reliable difference ($\chi^2=61.336$, $df=54$, $p = 0.230$) between the expected and observed frequencies in pre-test scores between the three conditions. The participants in both conditions seem to come from the same population, with a lack of statistical difference in prior knowledge across conditions. This means that inherent differences in the population would not be able to account for any differences found in the post-test, or the proportional learning gain within the session.

The results of the learning gain concerning the three software interfaces (u-ELM, AILM, and NLM) were compared using the Kruskal-Wallis Test for each learning session, which results are illustrated in Figure VII.3:

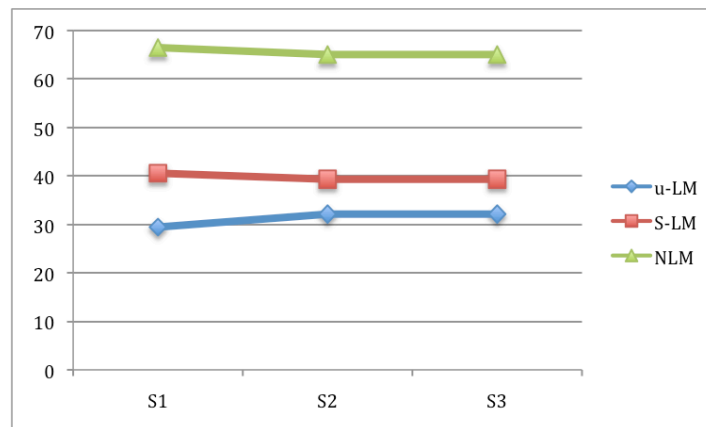


Figure VII.3: Average learning gain (%)
per interface per session

A statistically significant difference between the three different conditions was found for each learning session ($H_{S1}(2)=32.287$, $p<0.001$, $H_{S2}(2)=26.503$, $p<0.001$, $H_{S3}(2)=26.503$, $p<0.001$). H_0 is rejected for each session, and after performing a post-hoc Tukey multiple comparisons test, it appears that children learned statistically less knowledge, using the u-ELM condition than the NLM condition for each learning session ($q_{S1}=-0.080$, $p=0.001$, $q_{S2}=-0.084$, $p=0.001$, $q_{S3}=-0.128$, $p=0.001$). However, there is no significant difference in learning gains between the conditions AILM and NLM ($q_{S1}=-0.079$, $p=0.002$, $q_{S2}=-0.095$, $p=0.002$, $q_{S3}=-0.128$, $p=0.002$) or between the u-ELM and AILM conditions ($q_{S1}=-0.003$, $p=0.986$, $q_{S2}=-0.011$, $p=0.916$, $q_{S3}=-0.015$, $p=0.916$).

Number of LM inspection

A statistically reliable difference between the total number of LM inspections recorded for the different conditions was found, when considering all the learning sessions ($H(2)=20.912$, $p<0.001$) with a mean rank of inspections going from 92.87 for the u-ELM interface, and 99.42 for AILM, to 182.42 for NLM. H^0_3 is rejected, and it can be concluded that children statistically inspected more their learner model in the NLM condition than in the u-ELM or AILM one.

After performing a post-hoc Tukey multiple comparisons test, it appears that the NLM condition seems to lead children to inspect statistically more their LM than the u-ELM ($q_{S3}= 38.145$, $p<0.001$) or the AILM interface ($q_{S3}= 28.475$, $p<0.001$).

The results of the number of LM inspections performed on the three user-interfaces were compared using the Kruskal-Wallis Test for each learning session, which results are illustrated in Figure VII.4:

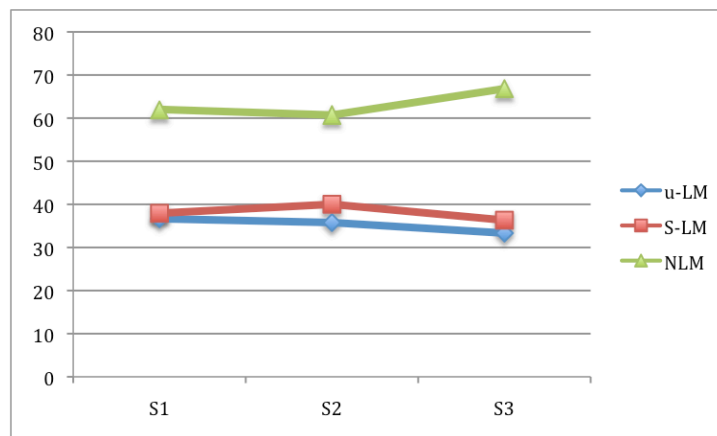


Figure VII.4: Mean of the number of LM inspection per interface per session

A statistically reliable difference between the number of LM inspections observed for the different conditions was found for the learning sessions ($H_{S1}(2)=17.818$, $p<0.001$, $H_{S2}(2)=15.707$, $p<0.001$, $H_{S3}(2)=30.386$, $p<0.001$) with a mean rank of inspections for the first session going from 36.67 for the u-ELM condition, and 37.93 for AILM, to 61.90 for NLM. H^0_3 is rejected for all three sessions, and it can be concluded that children statistically inspected more their learner model in the NLM condition than in the u-ELM or AILM one.

After performing a post-hoc Tukey multiple comparisons test, it appears that no statistical difference could be found between the conditions u-ELM and AILM for the each session ($q_{S1}= 0.015$, $p=0.976$; $q_{S2}= -0.024$, $p=0.783$; $q_{S3}= -0.031$, $p=0.605$). However, in every session, the NLM condition seems to lead children to inspect statistically more their LM than the u-ELM ($q_{S1}= 17.374$, $p<0.001$; $q_{S2}=15.638$, $p<0.001$; $q_{S3}=28.435$, $p<0.001$) or the AILM condition ($q_{S1}= 16.578$, $p<0.001$; $q_{S2}= 14.527$, $p<0.001$; $q_{S3}= 26.612$, $p<0.001$).

When considering the total number of inspections registered for the three sessions, a chi-square ‘goodness to fit’ test, was performed. The test failed to demonstrate a statistically reliable difference ($\chi^2=63.495$, $df=54$, $p = 0.239$) between the expected and observed frequencies in the total number of LM inspections between the younger and elder age-groups. The children involved in the study seem to want to inspect the model without any developmental difference to be seen between the seven to nine, and the nine to eleven years old participants.

The behaviour of children toward selection the activities to perform, when inspecting LM, was quantified according to the five categories defined in section 7.3, as illustrated in figure VII.5:

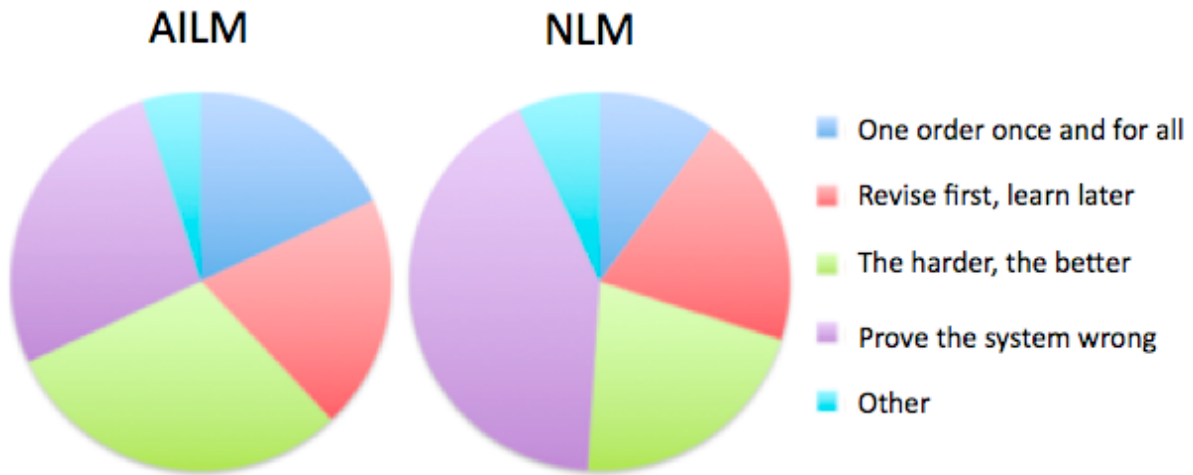


Figure VII.5: Children’s choice of activities between the AILM and NLM

When considering the categories 1 (“one order once and for all”), 2 (“revise first, learn later”), and 3 (“the harder, the better”) separately, children’s behaviour seems to be distributed in the same proportion in both interfaces.

Between the u-ELM and NLM the use of “medium” level feedback in the assessment seemed to lead children to considering his/her next action carefully, with comments such as “*Mmm, what should I do... Here ! on the table of 7, Multiplio doesn’t think I know it all. Perhaps I should revise the table and try it again...*”. This behaviour might be why children working on the user-editable model, without any system’s LM feedback, were found to interact less with the learner model than in the others.

In the AILM and NLM conditions, category 4 is represented with respectively 21% and 37% of children’s instances where they try to “prove the system wrong”. This behaviour, most commonly seen in the NLM interface, shows the greater level of interaction with the learner model, with children checking up the model for discrepancies when choosing an activity, and looking for the changes in the model once the activity is performed. This is in line with the precedent results on the number of LM inspections: as more children fall into category 4 in the NLM condition, the results of children’s LM inspection in this condition tend to be higher than in the other conditions.

Number of LM modification

A statistically reliable difference between the total number of LM modifications observed conditions, when considering all the learning sessions ($U=132$, $n_1=n_2=30$, $p<0.001$) with a mean rank of modifications going from 126.28 for the u-ELM interface, and 92.94 for AILM , to 180.82 for NLM. H_0^5 is rejected, and it can be concluded that children statistically altered more their learner model in the NLM condition than in the u-ELM.

The results of the number of LM modifications performed on the three user-interfaces were compared using the Mann-Withney U-test for each learning session, which results are illustrated in Figure VII.6:

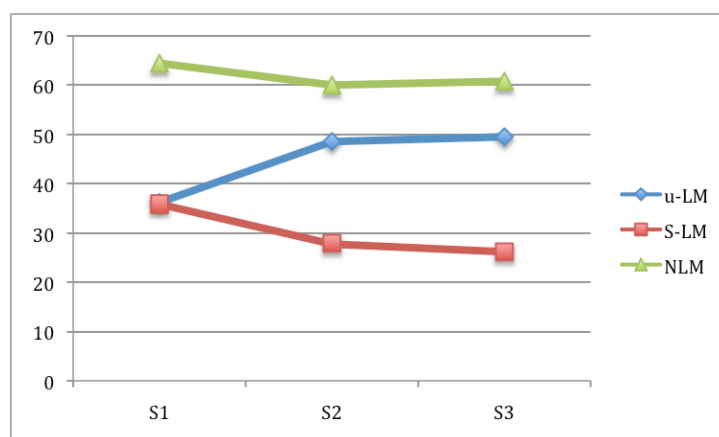


Figure VII.6: Mean of the number of LM Modification per interface per session

A statistically reliable difference between the number of LM modifications observed for the different conditions was found for the learning sessions ($H_{S1}=135$, $n1=n2=30$, $p<0.001$, $H_{S2}=142$, $n1=n2=30$, $p<0.001$) with a mean rank of modifications for the last session going from 49.52 for the u-ELM condition, to 60.82 for NLM. H^0_5 is rejected for all three sessions, and it can be concluded that children statistically made more modifications on their learner model in the NLM condition than in the u-ELM one.

Figure VII.7 shows the representation of children's behaviour when concerned with constructing the LM in collaboration with the system. Each time children were confronted with a discrepancy in the system's and child's view of the model in the NLM condition, their next action was quantified according to the four responses defined in section 7.3, as illustrated in figure VII.7:

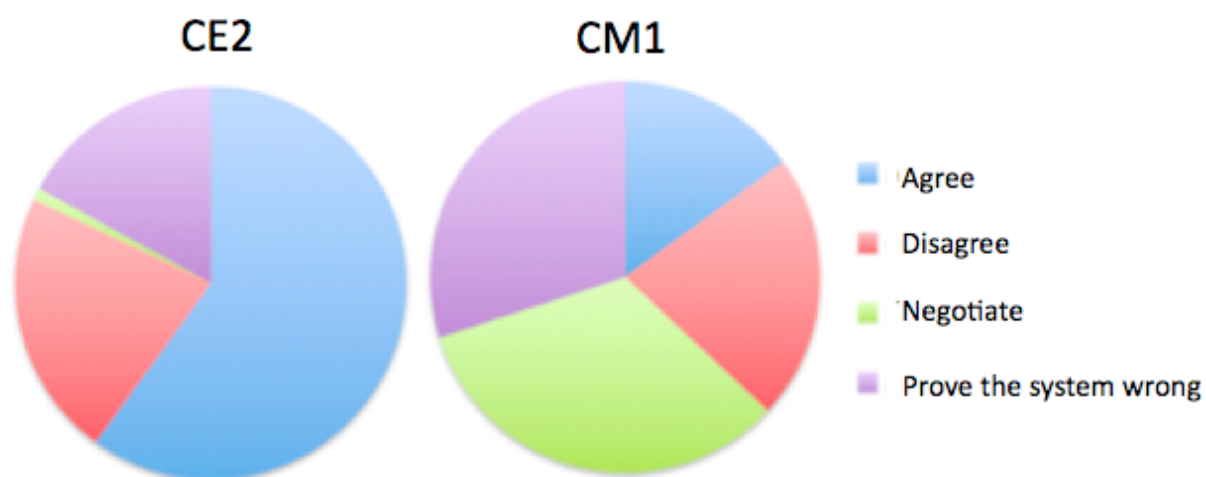


Figure VII.7: Children's negotiation of the LM per cohort

Children aged seven to nine tend to avoid negotiating with software on the model content, with only 1% of negotiation reported. A majority of the younger children (over 60% of the instances) prefer agreeing with the system's view. When they disagree strongly with the assessment, they chose to present a child's LM view that disagrees with the system's view (22% of the instances), or performed the activity in question in order to change the system's evaluation of their knowledge by proof (17%). The latter usually concerned strong disagreements on the assessment (from 'very bad' to 'very good' for example), and produced often emotionally charged verbal cues from the users, such as "No way!", or "What?!? You're wrong there, Multiplio, I know this table!"

Children from the elder group, however, engaged more in the LM construction activity, with 33% of the instances leading to a negotiation between user and system, and 31% of the instances where children performed the activity to change the system's view by proving their knowledge.

Emotional Input:

Number of Self-reports registered

The results of the number of self-reports of emotional state performed on the three user-interfaces were compared using the Kruskal-Wallis Test for each learning session, which results are illustrated in Figure VII.8:

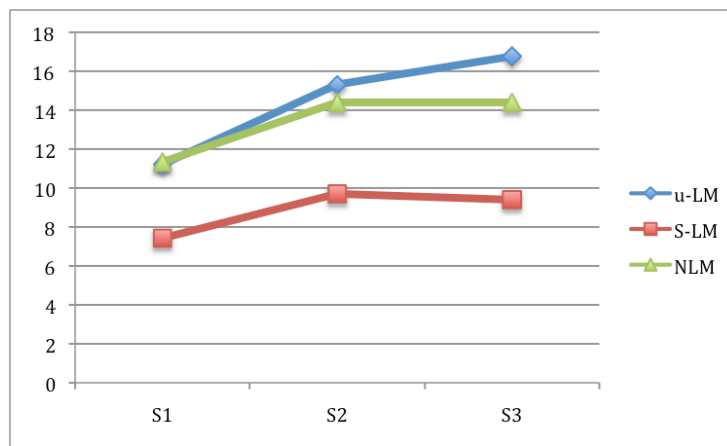


Figure VII.8: Mean of the number of self-reports per interface per session

A statistically reliable difference in the number of self-reports registered for a child working on the different conditions was found ($H_{S1}(2)=20.215$, $p<0.001$, $H_{S2}(2)=31.304$, $p<0.001$, $H_{S3}(2)=32.725$, $p<0.001$) with a mean rank of 11.215 during the first session (resp. 15.304 for the second, and 16.765 for the third) for the u-ELM interface, 7.425 during the first session (resp. 9.703 for the second, and 9.402 for the third) for AILM, and 11.341 the first session (resp. 14.398 for the second, and 14.398 for the third) for NLM. H^0_7 is rejected and it can be concluded that the children working the NLM registered a statistically reliable higher number of self-reputation than the children working on the AILM or u-ELM conditions.

After performing a post-hoc Tukey multiple comparisons test, it appears that during each learning session the AILM condition seems to lead children to report statistically less their emotional state than the NLM ($q_{S1}= 15.732$, $p<0.001$, $q_{S2}= 22.426$, $p<0.001$, $q_{S3}=23.004$, $p<0.001$) and the u-ELM condition ($q_{S1}= 17.262$, $p<0.001$, $q_{S2}= 20.208$, $p<0.001$, $q_{S3}=20.487$, $p<0.001$).

Representation of the emotions

A Chi-square ‘goodness to fit’ analysis, performed on the number of *Sorémo* self-reports registered for children across all OLM conditions revealed a statistically reliable difference between the two children age groups: Children aged seven to nine years seemed to statistically make more reports on their emotional state using the *Sorémo* tool than their elders for the first ($\chi^2=58.891$, $df=1$, $p<0.001$), second learning session ($\chi^2=57.625$, $df=1$, $p<0.001$). When considering the number of reports across the three learning sessions the same tendency can be observed ($\chi^2=60.841$, $df=1$, $p<0.001$). However, the results are not significant for the last learning session ($\chi^2=13.682$, $df=1$, $p=0.053$).

It is possible that the difference between groups is related to a developmental difference in how children choose to report emotional states using Soremo, with older children making less reports but using a complex set of emotions to grade each time. This would explain the data registered for the third session : if we consider the time spent between first and last use of software, children had the time to age and change behaviour. This could lead the older children from the CE2 cohort to act more like the CM1 ones.

Studies investigating children’s awareness of multiple emotions as part of an emotional state highlighted three kinds of representation: a unique emotion, a succession of unique emotions, or the presence of multiple emotions at once. Figure VII.9 shows the number of occurrences of emotional states reported as one of these three representations across all learning sessions and all conditions, distributed across the two children age groups.

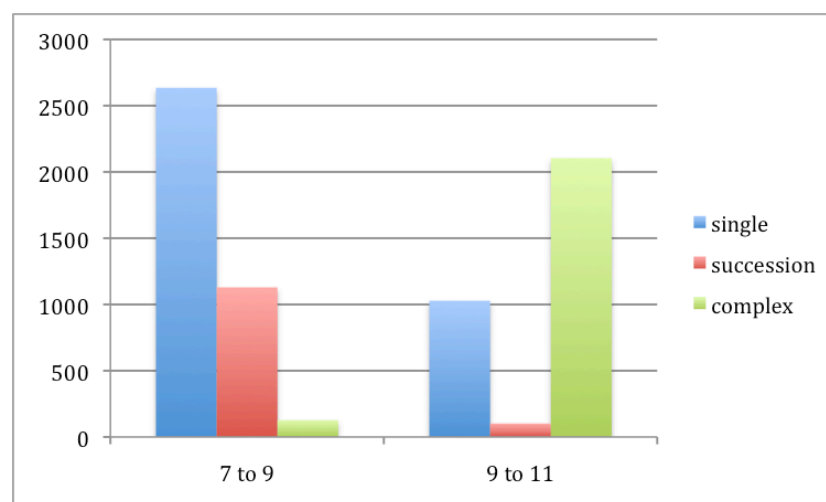


Figure VII.9: Repartition of the emotional states represented as single unique, succession of unique, or complex view of emotions

As illustrated in the figure, children aged seven to nine primarily express their emotions with a single rating of an emotion, and to a lesser extent, with a succession of reports containing a single rating of an emotion. On the contrary, children aged nine to eleven tend to represent an emotional state as a combination of complex emotions.

This tendency is in line with Harter & Buddin's (1987) and Stein et al's (2000) theories about children's development of the awareness of multiple emotions: Children aged seven to eight tend to view any emotional state as a unique emotion. Some of them can describe situations that will elicit two emotions, but only with one emotion following the other. By nine years of age, children's understanding of an emotional state includes a set of basic emotions and can they can think of emotions being provoked either successively or simultaneously. From these observations, an analysis was performed on the number of *Sorémo* self-reports registered for children across all OLM conditions, by considering the succession of several unique ratings of an emotion (without any other action taken on software between reports) as a single report. A Chi-square 'goodness to fit' test on this measure failed to reveal a statistically reliable difference between the expected and observed number of reports between two children age groups ($\chi^2=1.9745$, $df=1$, $p=0.1600$).

Distribution of the emotions reported

Figures VII.10 through VII.12 show how frequently each emotional state was scaled by children in each of the three user-interfaces (u-ELM, AILM, or NLM conditions) and distributed across the three learning sessions, according to the dimensions of arousal and pleasantness (green: >100; orange: >50; red: >10).

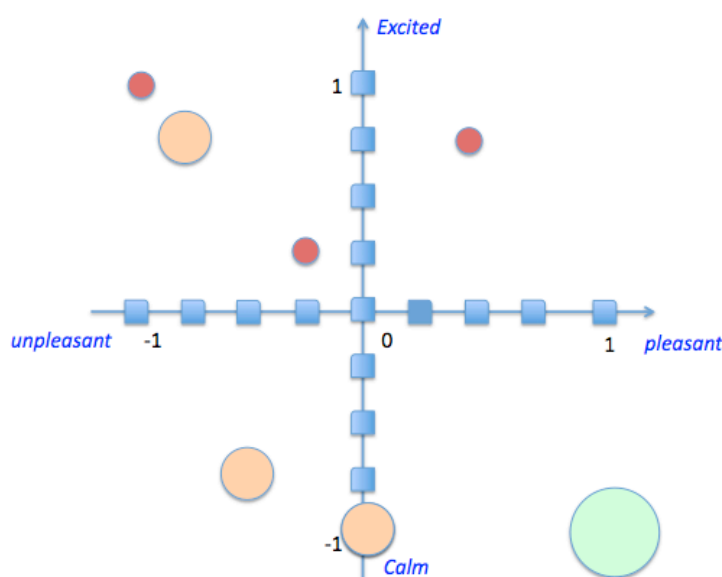


Figure VII.10: Grouping of the *Sorémo* emotions in terms of arousal and pleasantness during learning sessions for the AILM interface

Children working on the AILM condition seemed to use the *Sorémo* method sparingly, to express few emotional states: A strong tendency can be seen with the expression of emotions of low arousal and high level of pleasantness, such as ‘happy’. Few emotional states were reported as low on arousal and of neutral valence, such as “puzzled” or “thoughtful”. They were expressed in the middle of an activity, or when observing the LM content, and were preceded or followed by a few seconds of a lack of interaction on software. The user being in “deep thought” about the material proposed could explain this. Finally, some of the children reported on a few of emotions highly unpleasant and high in arousal, such as “disappointed” or “angry”. Such emotions were generally reported when *Multiplio* provided users with a feedback on their results the users disagreed with.

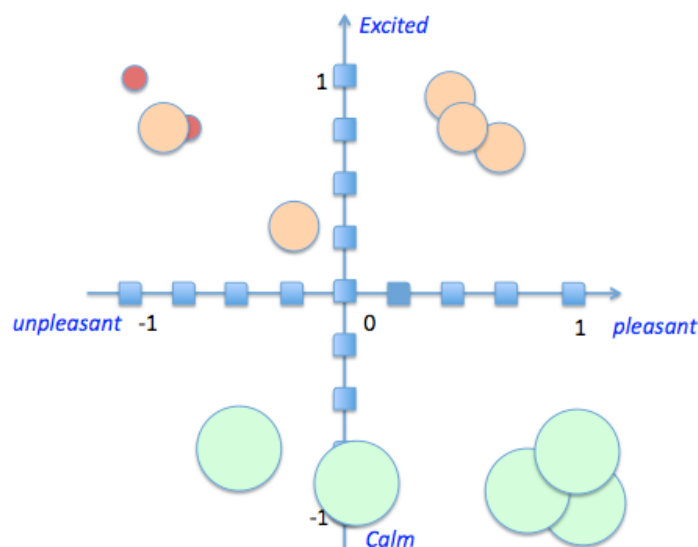


Figure VII.11: Grouping of the emotions in terms of arousal and pleasantness during learning sessions for the u-ELM interface

Children working on the u-ELM condition, however, reported emotions on the whole emotional space: A strong tendency can be seen once again with the expression of emotions of low arousal and high level of pleasantness, such as ‘happy’. The emotions of neutral valence and low level of arousal are also present in this sample, only with more instances. All emotional states of low level of arousal were reported at the beginning of the sessions, or between activities. Few reports express, similarly to the AILM condition, emotions of high arousal and negative valence, usually within an activity. Finally, a new spectrum of emotions appears here, emotions of high arousal and positive valence, such as “*captivated*” or “*satisfied*”. These emotions were reported once the user won an activity, before changing the content of the learner model on this activity.

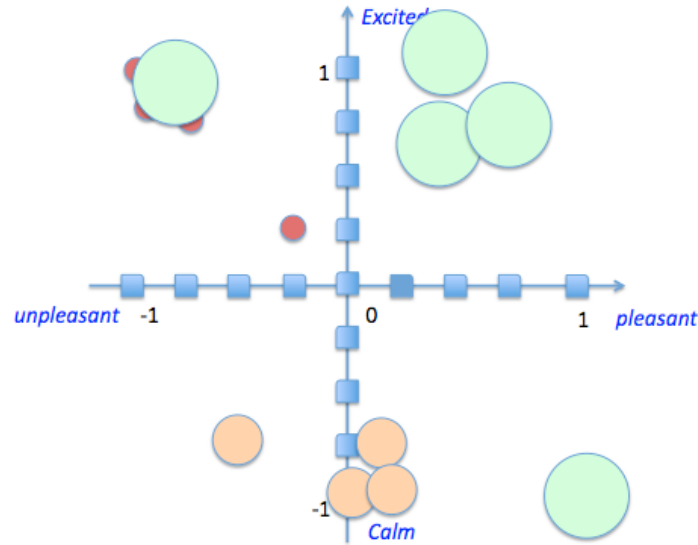


Figure VII.12: Grouping of the emotions in terms of arousal and pleasantness during learning sessions for the NLM interface

Finally, children working on the NLM condition reported emotions on the whole emotional space. However, unlike in the u-ELM condition where emotions of low arousal were favoured, children expressed their emotional states in the NLM condition as highly aroused: While emotions of low arousal and high level of pleasantness or of neutral valence are present, far less instances could be found when compared to the results of the u-ELM condition. In contrast, emotions of high arousal, both of negative and positive valence, were found. In addition to the situations prompting a report, described in the u-ELM and AILM descriptions, reports were given in this condition in relation to the edition of the learner model: after a negotiation process, or a test performed to prove one's opinion, emotions such as "satisfaction" were reported. Similarly, when discrepancies were spotted on the model between the system's and the child's view, vocal assertions were associated to reports of highly unpleasant and aroused emotions, such as "*indignation*", or "*disappointment*".

Research on the expression of emotions identifies two moments of elicitation: the emotions the user is currently feeling, and *anticipatory emotions* – emotions that users feel in anticipation of an event to occur. The reports were therefore categorized according to the moment of elicitation: emotions of high arousal, reported before feedback was given, or a negotiation process accomplished, were considered as anticipatory emotions.

Figure VII.13 shows the number of *Sorémo* reports made for each type of elicitation, in total across all learning session and user-interfaces, for the two age groups.

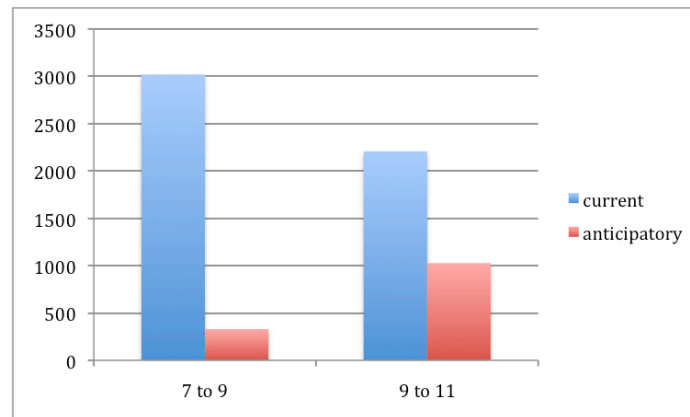


Figure VII.13: Repartition of the emotions scaled in anticipation or as currently experienced emotional state

Children aged seven to eleven seem to report emotions as they are felt, while the eldest groups tend to also report emotions they are feeling in anticipation of the next action on software, usually related to system's feedback. These results are in line with Bloom's (1998) and Harris' (1977) studies: the seven to nine years used the reporting tool to express emotions they were currently feeling, following an action on software that produced this emotional awareness. While few seven to nine years reported how they would feel in the future, the group of nine to eleven years also produced reports to that effect: they used *Sorémo* in anticipation of an emotion to come. Indeed, children usually reported an anticipatory emotional state, for example, before finishing a test with all the questions correctly answered. Expecting *Multiplio* to turn green and happy because of their success, the emotion reported were high in arousal and positive in valence, and an emotion such as "satisfaction" is to be expected after viewing feedback. Once the feedback was given, another report was filed, with the emotion anticipated at various levels of arousal, in this case "satisfaction".

Children's post-experiment Interview

Post-experiment interviews were conducted with each child participating in the study, as well as the two class teachers. Children were asked to present software features and how useful they were, and what they learned from software. They were also asked questions concerning software usability, as well as their willingness to continue using software the year after. The teachers were asked how children described their experience with software between sessions. They were also asked to point out the differences they had observed in children's understanding and manipulation of multiple concepts when learning in class between sessions: whether it changed their perception of mathematics in general, and multiplications in particular; how often they referred to software when doing multiplication exercises in class; and whether they believed software helped children learn this particular concept. All comments in this section are translated in English (all comments were made in French, as the study took place in a French classroom).

Children's interviews

Children seemed to refer to software goal and activities by the embodied-pedagogical agents used in each condition: *Multiplio* for system's view, and *MoiJe* for the user's view. It seemed to define the way children interacted with the OLM application: children took *Multiplio*'s view into consideration when choosing their next activity on software.

The presence of the system's view seemed to have made an impact on children's level of motivation to use software: Children working on the AILM and NLM conditions seemed motivated to use software during the three sessions, a majority of them enthusiast to use software again the year after. On the contrary, children working on the *user-editable learner model*, with no representation of the system's assessment, seemed to enjoy working on the interface the first time, but progressively consider the application less useful, or even "boring" in the following sessions.

Children working on the *inspectable learner model* found the display information about the system's assessment useful, especially to guide them in which activity to perform next. Some of the children expressed their choice of activities as trying to make the character "Multiplio" as happy as possible, as illustrated by this comment: "*Well, I wanted Multiplio all green, because he's happy when he's green. Here, you see, I did it, now Multiplio is green for my tables of 1, 2, 3, and 5!*"

Children working on the *negotiated learner model* seemed to use the information on the learner model provided to inform their next action, as illustrated by the following comments: "*You see, Multiplio is so cool, he knows who I am, and what I can do... When I go there [LM inspection screen], he tells me what I can do, and what I don't do well. Oh, here... my table of 6 is orange, I did some*

wrong yesterday [date of the last session]. I have to practice this table for next time, so I can show Multiplio... ”.

They seemed to navigate easily through the learner model content, and the activities to choose from. Here is a typical example of child’s description of software interface: *“This is the game. I like it. Here there’s all the things I can do... first, there’s the tables: table of 1, [...], and then Multiplio mixes all tables, that’s here [level 2]. And you know what, I looked here [level 3 and 4], and the numbers are HUGE, Multiplio gives you numbers with three and four digits... but I’m not good enough yet, I tried that [level 2], but I didn’t get it right, Multiplio was sad, and he told me I need to practice my multiplication tables again before I try it. [...] And here, you have all the tables, you can see them and practise before taking a test. [...]”*

One of the comments more recurrent was how useful the presentation of both views of the LM was: how it helped the children become more aware of their strengths and weaknesses in learning multiplications. When explaining the interface features, a majority of the children stopped a few minutes on the symbols representing the LM views of one particular concept, with comments such as *“[...] and here the game shows how good I am at the table of 3: I think I know it but I still do some wrong so I put Moije in orange. Multiplio agrees with me, see, it is orange too. [...]”, “What I like is to go where Moije is not like Multiplio, ‘cause sometimes Multiplio is wrong.”, or “Moije and Multiplio show you everything : Moije shows what I know I can do. And here, next to it, Multiplio shows me what he thinks. When they fight, I have to practice the table,... mhm... or I test this table to show that he [Multiplio] is wrong.”*

Teachers’ interviews

The teachers’ interviews revealed that children talked about *Multipliotest* and the activities they performed on software outside from the three learning sessions. A majority of children working on the AILM and NLM conditions seemed to dedicate their practice of multiplications to be able to “beat Multiplio”, and show progress in the next session on software.

The teachers reported numerous demands from children to access software at home between sessions to practice multiplications for the classroom tests, especially from the low-graders. Software was apparently found a useful tool to practice the tables, and help go back to the basics when the class learned multiplications of big numbers.

When children learned how to multiply big numbers, they often referred to the method presented by *Multiplio* in software. Similarly, when they received feedback on their classroom assessments in multiplications, they made comments such as *“Yay! Take that, Multiplio”*, or *“Multiplio would be green, I’m so good”*.

7.5 DISCUSSION

This study investigated children's use of different forms of OLM applications: inspectable, user-editable, or negotiated learner model. The results presented here give some evidence as to the utility of a negotiated learner model with graphically representing two views of the learner model (the system's assessment and learner's self-beliefs) to help children learn better, interact more with the learner model content in visualizing and editing the information provided.

Results from the experiment show that the children who worked on the NLM seemed to learn more from the pedagogical activities than the ones working on the inspectable or user-editable LM: they seem to learn more about multiplications when using software, as portrayed by their learning gains between the pre- and post- test scores.

Children working on the AILM and NLM interfaces were sensitive to the affective feedback given by the EPA representing the system's assessment (*Multiplio*) during an activity or in the representation given of the model. They reacted strongly, emotionally, to the under-estimation of their abilities by the system, and this influenced their next action in software: showing "*Multiplio*" that they were right about their assessment. The characters became alive, a representation of some sort of expert judge, or the teacher's voice, to whom they "had to prove" that their assessment was correct. Comments such as "*No way! I know my table of 2 by heart, you're wrong! I'm going to show you that you're wrong*" could be heard during software use.

The results on the number of inspection of the learner model and modification of the information also showed a strong tendency for children working on the negotiated interface to interact more with their learner model than the ones working on the other interfaces. The difference was more obvious from the second use of software: while the children working on the user-editable interface tended to go directly to the activities they knew how to perform without needing practice, the children presented with a representation of the system seemed to access this information again, looking for changes, and updates on the model information.

On the one hand a difference in the number of learner model inspection could be found between the two children age-groups (seven to nine and nine to eleven years of age). On the other hand, the youngest children were reported to modify statistically less their learner model. This difference is particularly salient in the negotiated interface, where the eldest children made more than twice the number of modifications the youngest children performed. It seems that while the seven to nine years were as keen as the nine to eleven years to go and inspect the model content, as an inspiration for future actions decision, they revealed to participate less in the join-construction of the learner model. Indeed, when confronted with a disagreement between their beliefs and the system's

assessment (for example the case illustrated in Figure VII.1), the majority of youngest children preferred to agree with the system's decision, rather than engaging in a negotiation process to prove their view. This behaviour might be explained by the way children conceive what the system's view represent. In the post-session interviews, it appeared that 'Multiplio', the system's representation of the LM through an EPA, was considered *alive* by children: a character capable of rational thought and expressing feelings. For most of the children involved in the study, *Multiplio* then represented what the teacher thinks of their achievements. Children aged seven to nine are not used to argue learning assessments provided by teachers, and therefore might have preferred a less 'confrontational' approach of disagreeing with the system. Indeed, when the system's representation of one skill was deemed overly under-estimated, they expressed vocally their astonishment or even affront, and made a point in choosing the activity to "prove the character wrong". The eldest children, when confronted to this situation, were keener on explaining their view and, when an agreement couldn't be made, go back to the activity to justify their assessment.

The three OLM interfaces seemed to produce different behaviours from children in the long term: all children first explored the interface, and then began working on the activities, with more or less interaction with their learner model. From the second session however, children, recalling which services the interface offered from the session a month before, did not carry on exploring, and two tendencies could be observed:

- Children working on the user-editable interface (modification of the student's view only, with no system's view available), went directly to choosing activities, and did not modify their learner model view much during the whole session.
- Children provided with a view of the system's assessment generally chose to access this information first and, then choose an activity. At the end of each activity, they went back to the learner model information to consider their next action.

In Bruce's theory of "free-flow" (2001), children learn better by being in control of their own representation of knowledge acquisition. From the study results, it seems that children favour a mixed level of control rather than full control over the system's representation to show how well they are doing, and reflect on their learning acquisition during the session. This is in line with Zapata-Riviera's study using a Negotiated learner model with Bayesian representations (2003), and Mabbot et Bull's study (2006), where it was suggested that users may prefer an OLM that offers less direct control, such as negotiation process. It may be related to the use of "play-props" defined as critical to enable "free-flow" in (Bruce, 2001): in the negotiated learner model, the representations of the student's and system's beliefs are directly comparable without extra cognitive load due to the mental representation of their own beliefs, present in an inspectable OLM. In the user-editable system, the only feature offered to be used as "play-props" is the character representing the child's representation of his/her

own skills. This character reacts as the user tells him/her to, but there is no confrontation, or feelings of satisfaction in case of success, or motivation to win in order to please the character. This reduces the level of “play-pretend” when compared to the negotiated learner model.

From the informal comments made by children after the experiment, in the closing intervention with the school, it seems that:

- Children working on the *inspectable learner model* found the display information about the system’s assessment useful, especially to guide them in which activity to perform next in order to “please Multiplio the most”. The only comments made considering their self-beliefs correspond to situations where the two assessments were opposite (concept not understood vs. concept fully grasped). In this case, children seemed to naturally want to “prove the system wrong”, by working on the activity corresponding to the concept in question, and hopefully do better. Vocal comments were added to this intervention, with some children showing disappointment of the disagreement discovery, and going back after the activities was performed to check the changes had been made. When the disagreement was less obvious, children generally seemed to pay less attention, and not to let it influence their next action.
- Children working on the *user-editable learner model*, with no representation of the system’s assessment, seemed to enjoy working on the interface the first time, but progressively consider the application less useful, or even “boring” in the following sessions. Some of them went back to the learner model display to update the changes in their knowledge that could happen by performing the activities, but most of the children stopped making modifications after a few activities.
- Children working on the *negotiated learner model* seemed to use the information on the learner model provided to inform their next action. After a period of software exploration, they went directly to the learner model presentation, and seemed to look for the concepts where a disagreement could be graphically seen between the child’s and system’s view of the learner model. Comments were made on the usefulness of “having what I know [child’s self-belief] right next to what Multiplio thinks [system’s assessment]”. A majority of children also liked to “talk with Multiplio”, and convince him that the child was right about the wrongness of the system’s assessment. After a reduction in the number of inspection/edition of the self-beliefs from the first session to the next, probably due to the end of the exploratory period, the number of children’s interactions with the learner model stayed high for the last two sessions.

One could argue that this study presents a potential confound between the amount of interaction children had with the LM and the type of interaction: the NLM interface requires more interaction from children with the model than the AILM or u-ELM conditions. Children might learn more by being offered more things to do, without consideration of the object of the actions. It could be that by initiating more attempts at the learning activity, or more reflection upon one particular learning process, children learn more. However the results presented here do not seem to agree with this position. Indeed, the children in the u-ELM condition were prompted to do as many actions in terms of manipulating the LM as the NLM ones. They could give their assessment on one concept, and choose to do an activity over if they so wished. The results in terms of learning gains are however statistically different: children using the NLM seemed to benefit more than the u-ELM ones on using software. The difference might not therefore be in the number of interactions but in the nature of the interactions themselves: comparing your view to the software's.

Sorémo was used to express emotional states of different levels of pleasantness and arousal (boredom, excitement, disappointment), just after the action on software that produce the emotion, or in anticipation of an action or a result to come. The emotional states reported by children in the study, using *Sorémo*, are in line with Bloom's (1998) and Harris' (1977) studies: the seven to nine years used the reporting tool to express emotions they were currently feeling, following an action on software that produced this emotional awareness. While few seven to nine years reported how they would feel in the future, the group of nine to eleven years also produced reports to that effect: they used *Sorémo* in anticipation of an emotion to come. These results are also in line with Harris et al's (1981) postulate that children around the age of six conceive emotions as an emotionally charged situation that provokes a visible emotional reaction. They furthermore rarely refer – unlike ten years olds and elder children – to the mental process that intervene between a situation and the emotional reaction that it provokes.

Children appeared to report more on their emotional states in the *user-editable* and *negotiated* conditions than the children working on the *inspectable* learner model. The emotional states reported appear to also differ between OLM applications. Children working on the *inspectable* software made few reports, with a majority of the reports low in the arousal level and highly pleasant. The children working on the *user-editable* interface, however, seemed to report emotional states with a lower degree of arousal than the ones working on the negotiated learner model. This seems in line with the level of verbal comments emitted during the learning sessions.

In the study, the number of reports given between the two age groups showed that the seven to nine years produced more reports than the nine to eleven years when considering all reports given. However, when removing the successive reports of one unique emotion without any change in software interaction, and counting them as one emotional state, the number of reports was equally distributed between the two participants' age groups. This shows that the younger children did not report more emotional states than the eldest group, but rather reported them differently.

The results confirm are in line with Harter & Buddin's (1987) and Stein et al's (2000) theories about children's development of the awareness of multiple emotions: Children aged seven to eight tend to view any emotional state more as a unique emotion. Some of them can describe situations that will elicit two emotions, but only with one emotion following the other. By nine years of age, children's understanding of an emotional state includes a set of basic emotions and can they can think of emotions being provoked either successively or simultaneously.

The use of the *Sorémo* tool, either in usability study or when embedded in pedagogical software, seemed to enable children to represent their emotional state according to the way they conceptualize it: either by a single report of a unique emotion, a succession of reports presenting a unique emotion, or a single report with multiple emotions scaled.

7.6 SUMMARY

This chapter is concerned with investigating how much children wish to gain access and control over the learner model content, and its benefits for learning. It aimed at answering the following research questions:

RQ7: How can children interact with their learner-model?

RQ8: What level of inspection and modification of the learner model content by the child-users is suitable for the child to learn efficiently ?

RQ9: What kind of learner model inspection/modification by users is most suitable to help children make use of the learner model to learn mode efficiently and gain more reflection on the pedagogical activity: full editing (user-editable learner model), no editing (inspectable learner model), or building the model with a user-system negotiation process (negotiated learner model)?

The experimental study presented in this chapter explores how the use of learner models composed of two different points of view, one representing the system's view and the other children's self-beliefs, can help children learn better with the OLM application by leading them to a greater reflection process on the learner model content and its application to the selection of pedagogical activities.

The results show that children do not interact in the same ways with an OLM application, according to its level of openness and student's control over the learning process. The children working on the negotiated learner model were found to benefit the most from using software, as their learning gain from before to after each learning session on software revealed to be higher than the one of children working on the user-editable or inspectable learner models. Two different representations of the learner model content were: one representing the children's self-beliefs, and the other the system's assessment of knowledge acquisitions. When integrated in a NLM, this has the potential to lead children to be more involved in the representation of what they know: children indeed use both representations to visually compare their views of how much a specific concept is grasped to the system's assessment. This leads them to engage in a negotiation process when a disagreement was found, enabling them to learn more from the sessions on software. The results indicate that children working on the NLM reported emotional states more intense than in the other conditions: the emotional states held a higher level of arousal and intensity, either as a highly positive (such as exhilaration) or negative (such as indignation) emotions. The reports correspond in parts to direct reactions to the emotional feedback provided by the EPAS, with children reacting more strongly to the pedagogical feedback provided in the NLM condition.

The next chapter concludes, outlining the thesis contributions to current research on affective OLM tutoring systems for young children, and presents future directions of research beyond the scope of this PhD work.

Chapter 8: Conclusion and Outcomes

8.1 THESIS SUMMARY

The aim of this PhD was to investigate if and how the use of open-learner modelling techniques and the inclusion of affective components in the design of intelligent learning environments could facilitate learning, and enhance software usability by increasing children's motivation to learn while using technology.

In chapter 1, current research (empirical findings, theories, and design practices) was reviewed, and the various issues concerning the building, evaluation, and use of affective OLM environments for child-users aged seven to eleven years were discussed. The research questions, to be answered in this Phd's work, were then defined as follow:

RQ1: What theories of emotion can be followed to produce a model of EPA's affective response in OLM software suitable for mathematical drill-and-practice applications?

RQ2: How can user-centred and participatory design techniques be used with children and their teachers to create affective components to be integrated in tutoring systems for children?

RQ3: What is an appropriate representation of a learner-model (LM) for a child to interact with?

RQ4: How can user-centred and participatory design techniques be used on children and their teachers to represent the Learning model components of OLM tutoring systems for children?

RQ5: Does opening the content of the learner model to the child-user facilitate learning and motivation?

RQ6: How can the use of pedagogical agents or the inclusion of affect in the design of OLM tutoring systems help children interact with their LM, facilitate learning and motivation?

RQ7: How can children interact with their learner-model?

RQ8: What level of inspection and modification of the learner model content by the child-users is suitable for the child to learn efficiently?

RQ9: What kind of learner model inspection/modification by users is most suitable to help children make use of the learner model to learn more efficiently and gain more reflection on the pedagogical activity: full editing (editable learner model), no editing (inspectable learner model), or building the model with a user-system negotiation process (negotiated learner model)?

The thesis' structure was defined as follow: In chapter 2, all aspects of the OLM software to be used for empirical investigations are presented and discussed. Chapters 3 and 4 then introduce usability and participatory-design studies that helped define, evaluate, and build two components of the applications: the affective component (presented in chapter 2 and evaluated in chapter 3), and the learner model component (introduced in chapter 2, then investigated and built in chapter 4). At last, chapters 5, 6, and 7 are empirical chapters investigating the possible use of mathematical drill-and-practice OLM applications by children aged seven to eleven:

- Study of the existent with traditional OLM representation tools embedded in the *Multipliotest 1.0* application (chapter 5),
- Investigation of how children can best interact with the learner model using affective embodied pedagogical agents as a model-interaction medium embedded in the *DividingQuest* software (chapter 6), and finally
- Investigation of how much children want to access, edit, and utilize information taken from the learner model content (chapter 7), using the results from chapters 5 and 6 in the design of the learner model content and the interaction mediums used in *Multipliotest 2.0*.

8.1.1 EXPERIMENTAL PLATFORMS

In chapter 2, the two OLM experimental platforms, support of all the studies performed on OLM software in their different user-interfaces and software version, are presented: the *DividingQuest* and *Multipliotest*.

First, the architectural model defined for both applications, is introduced. Inspired by Barnard & al's (1979), Brussilovsky's (1995), Dilenbourg (1996), and Van den Brande's (1993) work on open learning environments and ITS architecture, five components compose this model: the pedagogical model, the affective model, the learner model, the user-interface, and finally, the expert model. The later represents the brain of the application, linking together all other components and is in charge of transmitting information back and forth between the user-interface and the three other components.

This approach to OLM system architecture was coupled with the MVC model (Krasner & Pope, 1988; Dix et al, 2004) issued from research in Human-Computer Interface and Interaction. The separation between the 'model' and 'view' of the software enabled the implementation of experimental platforms able to support user-interfaces with different types of graphical tools, learner model representation, user feedback (affective, use of an EPA, textual feedback), and access/interaction with the model content. This facilitated the addition or modification of the different

information present in the OLM, communication with the model, and the degree of information gathered for one specific use of the platform. In the event that one experimental condition had to be added, only the expert model's behaviour will have to be extended and completed to this condition's specifics. Indeed, this module regulates the flow of information to the other components, and therefore reduces the cost of the platform reuse.

The use of this architecture for OLM software form a practical contribution of this PhD's work: the experimental platforms realized (*DividingQuest* and *Multipliotest*) could be easily reengineered, using the separation between form and content, in order to add or modify conditions of an experimental study. Indeed, extra components could be added, and extra user-interfaces could be formed to study different populations with a software-development less time-consuming and lower-cost than the traditional techniques, where modifications usually involve a change in software core.

Kok's (1991) and Bull & Kay's (2007) frameworks for OLM applications were found highly useful to scope the design space and help in the definition of the research issues related to OLM applications to be explored in this PhD's work. In particular, the SMIL© framework (Bull & Kay, 2007), usually used to compare the purposes and evaluation of completed OLM applications, was deemed useful as a design tool. It necessitates a strong investment to understand each module and consequences of sub-sections of said modules when related to the OLM space. However, it helps focus the scope of the researcher's application within this scope, and enable an easy check of the missing or conflicting features the application can bear within the design process, when considering such research questions.

To follow, the various aspects relating to Human-Computer Interface and Interaction are discussed: interaction styles, visual aspects (metaphors of learning and embodied pedagogical agents), cross-cultural use, participatory-design and usability of the applications.

Rationale as to the cross-cultural aspect of the study results is discussed, and led to the building and evaluation of pedagogical software, theoretical models, and usability tool usable and understood by French and English children aged seven to eleven. The two OLM software, the graphical representation of the emotions, and the development of the *Sorémo* method, therefore have included both French and English teachers and children, in order to produce software usable by children of both nationalities.

Two metaphors of learning, the *Traffic-Light system*, and the *Smiley Faces*, commonly used in French and English classrooms as pedagogical targets are introduced. These metaphors (section 2.2.2), understood by children when used in association with levels of learning, are at the centre of this PhD's work on affect: First a demonstration that a combination of the two metaphors to represent pedagogical targets was more beneficial for children's understanding of the targets and its

corresponding learning achievements than when using the metaphors separately (section 2.2.1 and Appendix B) was provided. The inclusion of colours following the Traffic-Light metaphor was then tested to help children understand affective states in digital characters better (section 2.2.2 and Appendix G). A separate study (section 2.2.2 and Appendix J), showed the benefits in terms of learning achievements, recognition of the emotional states digitalized, and motivation of adding colours to an affective EPA to express affect. From the positive results of these investigations, the Traffic-Light metaphor was added to facial/bodily expressions in EPAs to represent affect in every affective component of the applications: *DividingQuest*, *Sorémo*, and *Multipliotest 2.0*.

Finally, the different components of the OLM applications were introduced: pedagogical component, affective component, learner model component, and the user-interface component. In the pedagogical component, the pedagogical content of the *DividingQuest* and *Multipliotest*, are presented, as drill-and-practice mathematical applications, with an increasing level of difficulty. This component is common to all versions of the software, and user-interfaces.

The affective component details the two aspects that concern affect in the application: providing children with affective feedback on their learning activities and in the learner model representation; and offering a tool for children to express what they feel while working on software in the form of a self-reporting method.

The design of the model of affective feedback (A.M.E.R.) involved teachers from French and English primary school as design partners. The participatory design sessions led to the definition of a unique model of affective feedback to software interaction for the *DividingQuest* and *Multipliotest*. This model aims at testing the impact on learning of different theories on affect, children, and learning with regards to young users. During the definition of the feedback that would form the model, a negotiation process between teachers' in-school experiences and well-recognized theories (such as Kort et al's four quadrants (2001), the OCC model (Ortony et al, 1988), attribution theory, or Piaget's disequilibrium theory (1974), each introduced in section 1.3.2). The main characteristic of the model is the design choice of separating feedback according to teacher's evaluations of each child's mathematical abilities (section 2.3.2). The impact of the model in terms of children's engagement, motivation, and expression of emotional state were evaluated in the second part of chapter 3.

The results showed that the definition of emotional feedback in the model seemed to keep children engaged in the learning process and motivated. The participatory design process involving children as design partners detailed in chapter 3 led to the creation of a set of graphical animations using the *DividingQuest*'s EPA, the witch 'Séléna'. Each animation represented a possible feedback from the A.M.E.R. model, and was found easy to understand by children. Using children as design partners to design representations of affect understandable by this user age-group through the design

of affective digital character animations, helped reach an acceptable level of emotion recognition for such animations (average of 92% recognition).

The use of comics and mangas literature, as well as role-play in the participatory-design process was deemed essential to the success of the design activity. Practical experience results concerning the design process led to the production of some design recommendations that will be discussed in the contribution section of this conclusion. The benefits of using the model, by including the animations created in one interface of the *DividingQuest* software was evaluated in then chapter 6 in terms of learning gains.

The affective self-reporting tool, *Sorémo*, was developed using French and English children aged six to eleven as design informants and usability testers. The instrument appears to be understood, and usable for usability studies, and as self-report instrument to capture a child's current emotional state. After an exploratory and practice time of a couple of use of the instrument, children seem at ease with the emotional representations and the scaling procedure, being more decisive in their answers by using the middle scale only to show a variance in strength of the emotions felt rather than a lack of decision as to the presence of this emotion. *Sorémo* was then embedded in *Multipliotest* 2.0 to allow self-reports of children's emotional state, and tested in the investigation in chapter 7.

A development difference in the way children view affective states has been highlighted by the experimental studies of chapter 7: From the rating behaviours of the participants in order to express their feelings, it can be seen that some children aged seven to nine still view any emotional state as a unique emotion. On the contrary, the emotional states were perceived as a complex set of basic emotions by all participants aged ten to eleven. This is in accord with Izard et al's (1993) observations on young children's representation of emotional concepts for five to eight years olds, and Harter's (1977) theories and observations on children's development of mixed feelings understanding and expression. *Sorémo* is a fully operational self-reporting method, which has proven to be usable both in usability studies or when embedded in educational software. This second practical contribution could be used in the future to study users' emotional states reports in different learning contexts, for different population types, and users' age groups. Future research could include investigations as to users' understanding and manipulation of the tool when considering children of a different nationality and age groups, using the method either in usability studies or within OLM applications.

The learner model component of the OLMs, section 2.4.3, details all aspects relation to the model content: what information is included within the model, how it is edited at the initialization of software use by children, and modified while children play. The visualization and interaction mediums with the model content are rationalized, designed and illustrated in chapter 4. First, a study of children's favoured and most understood visualization techniques concerning learner model content guided the design of the learner model in the *DividingQuest* and *Multipliotest*. The 'lectures' and

‘related concepts’ views seem to be better recognized and preferred by children this age. Under those views ‘skill-meters’ and ‘boxes’, when utilized as learner model representation tools, seem to be strongly preferred and facilitate a better understanding of the underlying model. From those results, three sets of learner model representation were formed: in *Multipliotest 1.0*, traditional ‘skill-meters’ and ‘boxes’ were embedded in the ‘lecture’ and ‘related concepts’ views. In the *DividingQuest*, the skill-meters were still present in one interface as pedagogical targets designed using the traffic-light system, and EPAs were used in the two other interfaces as interaction medium with the model, in order to enhance children’s understanding of the model content and its uses. Indeed, as the research questions to be answered were related to the impact of using EPAs to interact with the model, only one interface held a ‘standard’ representation of the learner model, using the learner model view and tool best understood from the study in chapter 4. In *Multipliotest 2.0*, two embodied pedagogical agents, representing the system’s model and the child’s self-beliefs about the learner model, were integrated in the ‘boxes’ tools, for each concept to grasp, therefore allowing visual comparison of the two beliefs about the model content. The understanding and use of the different learner model representation were then tested in the experimental chapters (chapter 5, 6, and 7), as defined in the user-interface component (section 2.4.4).

8.1.2 EXPERIMENTAL RESULTS

Chapters 5 to 7 use software components and applications presented earlier, and designed specifically for French and English children aged seven to eleven years. They aim at investigating whether, and in what ways, children could benefit from having access to the content of their learner model for better learning, and motivation to learn, in affective OLM software.

In chapter 5, children’s willingness to use the learner model content is evaluated, and the benefits of an OLM when compared to software with the learner model hidden from the user, is investigated. Results show that pedagogical application that hides the learner content to the user (CLM) is less beneficial for children, especially low-skilled children, than software opening the learner content to the child-users (OLM), in terms of learning gain, and motivation to learn using technology. Children seemed to wish to access, use, and edit the learner model information. However, the access to the learner model information seemed to be used by low-skilled children only, with few children referring to the model for guidance as a systematic approach. The main concern when using the learner model seemed to revolve around the means to interact with the learner model and understand how grateful its use could be for learning.

From these observations and results, chapter 6 proposes a new way for children to interact with the learner model content. It aims at helping them better understand the content of the learner model and how it could be applied to gain more understanding of their achievements, influencing positively their planning of drills to try during the session: using affective embodied pedagogical agents as an interaction medium in the model. Chapter 6 therefore introduces a study investigating the potential of using EPAs portraying different emotional states to act as an interaction medium with the model, when embedded in the actual design of the learner model representations.

The results show that using the A.M.E.R. model of emotion, embedded in the software's EPA, helped children be more motivated in learning using software and making use of the learner model information provided when compared to when the EPA (or button screen) uses no emotional feedback. They seemed receptive to the emotions portrayed, and used them to guide their actions on software more efficiently. Results from children's interviews on their use of the learner model revealed a wish for more control over the display of the information, and a discussion to take place between the child-user and the system in order to build the model content more accurately.

In chapter 7, a final exploratory study was presented investigating the ways in which children are willing and able to use the information presented to them via affective EPAs to gain some understanding of their knowledge level, and guide their interaction with pedagogical applications by challenging or negotiating with the system's view of their knowledge level state.

The results show that children do not interact in the same ways with an OLM application, according to its level of openness and student's control over the learning process. The children working on the negotiated learner model appeared to learn more while using software, as their learning gain from before to after each learning session on software was shown to be higher than the one of children working on the editable or inspectable learner models.

The jointly built learner model in the negotiated version of software seemed to lead to greater interest from children to inspect and edit their learner model. The negotiation process that took place between the children and the system included children ranking their success and arguing with software's view via Likert scales, and redoing a test to support an argument. It also seems that the visual presentation of both their self-beliefs and the system's view of their acquisition level per skills impacted how children chose their activities in software. Some behaviours were observed: proving to "Multiplio" that the child is right by performing an activity s/he deemed under-evaluated, or perusing the model for guidance as to the activity to choose and engaging into the ones defined as in need for a deeper understanding of the concept.

Children found the affective self-reporting tool developed in this thesis and embedded in the three software interfaces, *Sorémo*, interesting and useful. Participants working on the editable and

negotiated learner model interfaces seemed to produce more emotional feedback than the ones working on the inspectable learner model: the children used more the self-reporting tool to express their feelings at different moments of the interaction. The tool was used to express different kinds of emotional states: *boredom*, *happiness* and *excitement*, *disappointment*... Children appeared to use the tool at two moments of interaction: *after an action* eliciting an emotional response, or *in anticipation of a result* in an activity. When the self-reports were used to express negative emotions such as frustration or disappointment, some children had the unfounded impression that the system took their feelings into account, and adapted the difficulty of the activity, or the feedback, in consequence. This behaviour, and the implications for learning of implementing such behaviour from the agents, could be investigated further, using the tool in different learning environments. Due to children's interest and actual use of the tool, their reports of emotional states could possibly be used in the future to inform other actors of the learning process on the children's emotional development (such as teachers, peers, or parents), or be integrated as part of the learner model component in the decision of the feedback given, as portrayed in the future directions section of this chapter.

The contributions of this thesis, described in the next section, should help give evidence of which theories of emotions better apply to children aged seven to eleven working on OLM applications, how children can, want, and effectively use learner model components according to its representation, content, and method of interaction, and therefore help in the design of future affective OLM educational applications for primary school children.

8.2 CONTRIBUTIONS

The contributions made to research from the results observed in the PhD will be presented in this section, according to four categories:

- ***Experimental Platforms***: the OLM applications made available from study use (*DividingQuest*, *Multipliotest 1.0*, *1.5* and *2.0*), and a affective self-reporting tool (*Sorémo*), and their architectural model.
- ***Building affective OLM applications***: the experience results from working with children and teachers as design partners, design informants, and usability tester to build different components of an affective OLM software and an affective self-reporting tool;
- ***Learning and Emotion theories***: the results of applying different theories of emotion for learning, and child emotional development, for OLM applications used by children from seven to eleven years of age, and its impacts for future theory development.
- ***Why, when and how to use OLM applications with children***: the results of studies investigating the use of affective OLM applications by concrete-operational children for better mathematical learning, increasing motivation, how and when children use and interact with information from their learner model, and the evolution of children's emotional state while playing on software.

8.2.1 EXPERIMENTAL PLATFORMS

Two OLM applications were developed during this PhD: new components were added to the *DividingQuest* and *Multipliotest* was created as an OLM application, under several versions, answering various research questions). They are all fully functional for class deployment, and accessible locally or in remote learning settings. Each piece of software includes several software interfaces, which enable different research questions to be addressed: in any experiment, a child was subjected to a particular interface, which shared similarities to the other interfaces tested, but also differed in terms of inspection/editing/visualization of the learner model component. Children and teachers participating in the various experiments expressed their willingness to use such systems in the future, and access to *Multipliotest 2.0* under the Negotiated-Learner Model form was given to the school participating in the final experiment for the 2009-2010 CE1 and CE2 classes to use during the year. It was consequently used for the whole year, children accessing software during classroom-based learning session, as well as at home, as homework or by personal initiative.

Part of the contributions linked to the creation of the OLM applications resides in the extension of the SMIL[©] theoretical framework, and the creation of evaluation metrics to inspect user's actions. When the SMIL[©] framework was originally created (Bull & Kay, 2005), the use of affect in ITS was not as pronounced as it currently is. Therefore no provision was given to the possible use of affective information, either as pedagogical feedback in the application, or as input data from the learner himself through intrusive or non-intrusive methods. Two boxes were therefore added in the framework to compare the applications presented in this PhD's work, in two of the elements of the LM: element 1 (extent of the model accessibility) and element 7 (presentation). When considering the literature on experimental evaluations of ITS and OLM systems, few mentions of the analysis of student's behaviour while playing on software could be found. As the experimentations of chapter 5, 6, and 7 were investigating children's use of the LM component, some metrics were needed to follow children's actions on software, a system to analyse children's patterns of interaction with the learner model was introduced, and another studying children's learning strategies according to their choices of learning activities, with a rationale given through qualitative analysis by experts.

8.2.2 LEARNING IN MATHEMATICAL OLM DRILL-AND-PRACTICE APPLICATIONS

The work presented in this thesis is scoped to children users aged seven to eleven years old, drill-and-practice mathematical OLM applications deployed directly in classrooms. In this context, learning has been defined as an “active, constructive, cumulative and self-regulated process in which the learner play a critical role” (cf. section 1.2.3).

When considering learning as such a process, results of children's manipulation of OLM applications seem to show inspectable and, to a greater extent, negotiated learner model applications as beneficial for children this age to foster learning. Indeed, children, playing a critical role in the good functioning of ILM and NLM applications, can learn *actively* due to their involvement in the representation of their knowledge base, the presence of instantaneous feedback, and the choice given to children over their sequence of activities on software. The learning within software is *constructed* through the child's representation and manipulation of the learner model content provided, which helps them link together the concepts to learn, and gives them a better understanding of their own achievements and limitations. Due to the nature of the mathematical applications investigated in this study, the *cumulative* properties of the learning process are enhanced, in the functionalities offered by the drill-and-practice application: children gain knowledge by repeating a process of testing their knowledge on specific concepts, one at a time, with drill identical in structure and pedagogical goals,

if not content. Learning in OLM applications is *self-regulated* by users due to their own choice in activities to choose, and takes into account their past and current performances to increase their learning capabilities, following a vygotskian learning strategy, based on the zone of proximal development (ZPD).

The concept of OLM applications was found as a potential support for the development of Vygotsky's "Zone of Proximal Development" theory (Vygotsky, 1977). Intelligent tutoring systems, using the information included in the learner model component, can easily use the zone of proximal development framework to propose individual instruction to children. OLM applications go a step further in the expression of this model, by allowing children to visually inspect and edit elements of the ZPD. Indeed, changing the content of the pedagogical information relating to the learner model modifies the activities given to the children, and therefore can increase or decrease the level of difficulty they have to face within pedagogical activities. The ZPD can be clearly see through the decomposition of the concepts to acquire in levels of difficulty, and the use of enabling/disabling features to go from one level of difficulty to another.

One concern in the OLM community is that by giving a total control to the learners over the LM content, the information of the LM corresponding to the ZPD could be set to values that are not in line with the child's knowledge, and potentially harm the child's learning process. In the experiment presented in chapter 7, results showed that some of the children working on the editable learner model ended up with learner model contents that did not reflect their achievements on software or achievements in class between software use. The user of the inspectable and, more particularly, negotiated learner model applications showed that a majority of children did reflect on the learner model content and edition, and learned from the LM information provided. This suggests that showing some indications of how the ZPD is implemented in system, in addition to the system actually using the ZPD to produce the drills, might be highly beneficial for the learning of mathematical concepts such as multiplications or divisions.

The results as to the benefits for learning of OLM applications and the inclusion of affect in EPAs as interaction medium are confined to the use of mathematical drill-and-practice applications. The experiments presented in this PhD are using mathematical concepts that were introduced in classroom at the time of software use. The children had learned the theory behind the multiplications and divisions concepts in class, and practiced the calculation of such operations in various paper-based tests. The drills within the OLM applications were similar in structure to the paper-based examinations, and therefore did not produce concepts or methods that were unknown to children prior to the experimentations. This is in line with the nature of the drill-and-practice applications, and focuses on the practice of skills and skills acquisition. The outcome of using affective OLM applications on learning environments with different learning practices are unknown and should be

investigated further. However, from the PhD's experimental results, a postulate was formed as to the good fit of the drill-and-practice applications with the OLM components: as discussed by Piaget (Piaget, 1970), such mathematical concepts rely on the learner's understanding of classification, reversibility, and conservation. The concepts learned here were the initial learning blocs that one would build upon to learn more complex multiplications and division concepts. The underlying structure of the learner model therefore followed the natural understanding of increasing difficulty of the activities: First, you have to learn the individual blocs, for example the multiplication tables. Only when you know all of them can you move further to long multiplications, etc... It is our belief that the results portrayed here would be similar for applications manipulating concepts with a similar clear structure, such as the apprenticeship of reading (beginning with the alphabet, then the sounds, syllables, words, and finally sentences), or foreign languages (with lower concepts including pronunciation, alphabets, and vocabulary, before moving on to sentences structures etc..).

8.2.3 APPLICATION OF THEORIES OF EMOTIONS FOR CHILDREN'S LEARNING USING TECHNOLOGY

In this PhD's work, emotion recognition, understanding, and expression were seen as an evolutionary process, throughout children's cognitive, cultural, and social development as portrayed in Ekman's discrete emotion approach (Ekman & Friesen, 1979), and extended by biologists, psychologists and research scientists such as Darwin (1998), Izard et al (1993), or Pons et al (2003). Three aspects belonging to Salovey et al's (2000) emotional intelligence concept were addressed in regards to the use of pedagogical activities by six to eleven years old: recognizing and understanding your own emotions; expressing your emotions to others; and managing one's changes in emotional state after elicitation.

As children recognize, understand, and respond to emotional situations displayed by software, they go through an emotion elicitation process themselves. In this PhD, an affective self-reporting tool was designed, and embedded into pedagogical software. This tool, *Sorémo*, enabled children to express their emotional state when playing educational games, and represent the emotional state of a digital character in usability studies.

The usability studies performed during the creation of the *Sorémo* method, on card-sorting activities (chapter 2), and the interactive animations of the *DividingQuest's* emotional feedback (chapter 3), are in line with Hall et al's (2007) results: children aged six to eleven are able to recognize a set of emotional states when portrayed by digital characters.

The understanding of said emotional states appeared to increase when the design included objects that help to put the emotion in context, and when the Traffic-Light metaphor was embedded in the design. While the valence of the emotions seems accurately recognized by a majority, individual differences in recognition of emotional states within one polarity (positive or negative emotion) have to be taken into account in the final results. The redesign of the initial *DividingQuest* animations (Girard, 2006), using children as participatory-design partners, helped in the definition of a set of emotional states deemed usable and easily recognizable by French and English children aged six to eleven.

Children's understanding of basic and complex emotions was tested in usability studies and when confronted with educational software components producing emotional feedback. The studies included card-sorting activities; descriptions of emotional situations in the form of a text, a drawing, or an oral explanation; and the use of Sorémo, the affective self-reporting tool developed in this thesis.

The results confirm Harter & Buddin's (1987) and Stein et al's (2000) theories about children's development of the awareness of multiple emotions: children aged seven to eight tend to view any emotional state as a unique emotion. Some of them can describe situations that will elicit two emotions, but only with one emotion following the other. By nine years of age, children's understanding of an emotional state includes a set of basic emotions and can they can think of emotions being provoked either successively or simultaneously.

This behaviour was observed in chapters 2 and 3, when developing and testing the *Sorémo* self-reporting tool in usability studies, and building the set of emotional responses with card-sorting and emotion recognition exercises: In line with the observations of Saarni's study (1988), the majority of children from the youngest age group (seven to nine years) reported emotional states of positive valence only; negative valence only; and in a smaller proportion 'mixed feelings' (some of the positive emotions were scaled, as well as some of the negative emotions).

Similarly, the results of chapter 7's experiment, when using *Sorémo* to report on their emotional state, show an similar behaviour: The seven to nine years submitted more reports than the nine to eleven years when considering all reports given. However, when removing the successive reports of one unique emotion without any change in software interaction, and counting them as one emotional state, the number of reports was equally distributed between the two participants' age groups.

The use of the Sorémo tool, either in usability study or when embedded in pedagogical software, seemed to enable children to represent their emotional state according to the way they conceptualize it: either by a single report of a unique emotion, a succession of reports presenting a unique emotion, or a single report with multiple emotions scaled. The method might therefore be more

adapted for children than the ones proposing a unique representation of an emotion, such as SAM (Sander & Scherer, 2009), or the AffectButton (Broekens & Brinkman, 2009). Indeed, children viewing their emotional state as mixed-feelings might have troubles using a tool that requires choosing one single emotion to identify with. These results can also be explained by the developmental difference in children's understanding of complex emotions and mixed feelings, as discussed earlier.

The study of theories of emotions relating to learning, as well as the input from participatory-design teachers, helped produce an **Affective Model of EPA's Responses** to software interaction (the A.M.E.R. model). This model tests the use of the emotions considered in Kort et al's (2001) Four Quadrant Model of emotions, as well as the control-value theory of achievement emotion (Pekrun, 2000), and the use of teacher's experience to facilitate emotional scaffolding (Williams et al, 2006) during software interaction, using the OCC model (Ortony, Clore and Collins, 1988) way of conceptualizing the emotions – in relation to objects and context – for the expression of affective feedback that highlights the pedagogical actions. The model, when tested against a lack of affective feedback in chapter 6, was proved beneficial in terms of learning gain, by virtue of increasing in child's motivation to use pedagogical software, and increasing the rate of learner model inspection.

All children participating to the studies in chapters 6 and 7 were not at the same stage of the development of emotional recognition. Therefore, the affective feedback of the A.M.E.R. model, given to children in software, was defined as a set of single emotional states, that could be understood by all. While this facilitated the recognition of the feedback by participants of all age, a research question can be raised: What is the real implication for instructional OLM software of some children understanding unique emotions only, when software proposes several in combination, or in close temporal proximity? What would be the impact on learning of a division in the child's understanding of affective information, pedagogical feedback for example, due to the child's awareness of multiple emotions? In this thesis, affective feedback was successfully embedded in OLM software to help the understanding of, and interaction with the learner model component. While the emotions chosen as feedback for the LM information were kept simple to understand, and separated in unique emotions, what would be the impact of multiple emotions on young children's manipulation of the model, their understanding, and therefore learning?

The results of use of Sorémo in chapter 7 show that the method was used to express emotional states of different levels of pleasantness and arousal (boredom, excitement, disappointment), just after the action on software that produce the emotion, or in anticipation of an action or a result to come. The emotional states reported by children in the study, using Sorémo, are in line with Bloom's (1998) and Harris' (1977) studies: the seven to nine years used the reporting tool to express emotions they were currently feeling, following an action on software that produced this emotional awareness. While few seven to nine years reported how they would feel in the future, the group of nine to eleven years also

produced reports to that effect: they used Sorémo in anticipation of an emotion to come (cf. section 7.4). These results are also in line with Harris et al's (1981) postulate that children around the age of six conceive emotions as an emotionally charged situation that provokes a visible emotional reaction. Unlike ten years olds and elder children, they rarely refer to the mental processes that intervene between a situation and the emotional reaction that it provokes.

As children expressed an interest in using Sorémo at various moments of interaction, it should be possible to draw a map of evolutionary emotional states of a child throughout software use. This point out interesting research questions, that were not investigated in this PhD's work: could the reports influence the emotional feedback given by EPAs, by taking into account an evolutionary emotional state component in addition to the pedagogical strategy implemented? What would be the impacts of making evolutionary affective feedback on children's representation of the characters, their trust in the OLM, and possibility of negotiating with the learner model content, as well as children's learning achievements? Could the reports be useful when presented in a synthetic form to other partners of the learning process such as the child's teacher, peers, or parents?

8.2.4 PARTICIPATORY-DESIGN OF AFFECTIVE TECHNOLOGIES WITH CHILDREN AS DESIGN PARTNERS

The participatory-design techniques used in this PhD, for the creation of software components, emotional representations, and pedagogical content, seemed to lead to the design of easily usable systems, motivating and fun to use by children, and deployable in classrooms for long-term use. The work performed with teachers led to the creation of an affective model of EPA's emotional responses to software interaction, from the combination of theories on learning, and teacher's experiences in classroom settings. The participatory-design engaging children as design partners for the definition of digital animations representing a character's different emotional states helped to design affective expressions believable and understandable by child-users aged seven to eleven.

During the process, two techniques were added to the techniques traditionally used: the use of comics as a mean of expression of ideas; and the inclusion of actors in the design partners to help children understand better the affective concepts and how they relate to human emotions. Those two techniques were extensively used in the design process, and seemed to help greatly in children's understanding of the concept, and motivation in developing such animations.

This method appears promising in terms of designing affecting components for educational software with children as design partners. One can wonder whether the design object – EPAs to be

embedded in a pedagogical game – could have impacted how well the result fit with the choice of adult design partners – comic designers and actors –, and whether this can be extended to other uses of affective design. The results of the application of this method for the construction of the EPAs seem encouraging on that aspect. It is true that the use of comic designers as resource for children's design of emotional display suited very well a gaming and fun concept. However, the children were not involved in the process of building the educational application. The animations, product of the design, were introduced to the children as a tool to be used in a pedagogical context, via the use of the scenarios of emotional display. They were not presented as part of a specific application or for a specific pedagogical scenario.

The main drawback of this method is the cost of including actors and comic authors into the process. However, the cost of design might be reduced by integrating the knowledge of such participants into the design team. Actually, most design teams have a number of people involved from different disciplines: teachers, academics in different domains of expertise, designers,... It would be interesting to investigate how the use of a couple of experts in comic design and drama could be added to the design team in order to obtain similar results, with only children as specific design partners.

The sessions of participatory-design with children and teachers holding different roles in the design, led to the definition of the following design recommendations for the design of pedagogical and affective components for OLM applications:

- *Use a diversity of experts as design partners: include comic designers and actors for affective products.*
- *Use Facial Action Units to manipulate expressions and emotional concepts*
- *Use objects to help put the emotions in context*
- *Vary the activities to help design the whole emotional space*
- *Integrate the design study in a class project*
- *Ensure that the understanding of the emotional concepts is the same for all design partners.*
- *Provide teachers with applied, descriptive examples of representation of the components to be designed, how they are linked together, and how they will be used in context in software.*
- *Provide high-level prototypes to teachers to represent software improvements, but also low-level prototypes material for them to sketch out ideas.*

Some assumptions were made about the design participants and the environment in which the final products would be deployed and tested: it was expected that the children end-users would have normal corrected vision and be able to use a mouse and keyboard as software interface input technology. They would also be expected to be in the ‘normal’ range of emotional development, able to recognize and express a set of basic emotions. While the redundancy of affective information in software (facial/bodily expression and colours) enabled coloured-blind children to use software, they would finally be expected to differentiate between the three colours of the traffic-light system metaphor: green, orange, and red; either by visualizing the colour, or differentiating their intensity.

8.2.5 FIELD EVALUATIONS OF PEDAGOGICAL APPLICATIONS IN CLASSROOM SETTINGS

When designing and executing empirical studies and experiments, it is essential to take into account the interaction characteristics of the technologies used and how they can affect end users [2, 3, 18]. It is even more crucial when the purpose of the study reaches beyond interactional usability concerns: e.g. when evaluating pedagogical applications, the most common factor within the study is related to learning, either as a measure of performance during one particular task, or learning gains related to software use over a lengthy time period. The issue concerns whether learning gain or lack relates to use of the technology, vagaries of the empirical study set-up context and pragmatics rather than the pedagogical paradigm. It is therefore important to gather case studies and form guidelines or recommendations to reduce usability issues and technical difficulties, and consequently plan studies that have better opportunities to actually investigate specific research questions without being hindered by usability issues.

Decades of researchers in Interaction Design for Children and Educational Technology have performed evaluation studies of various kinds of learners, using various technologies. Methods have been created to design, and evaluate products with regard to the human-computer interaction factors (Carroll, 2000; Dix et al, 2004; Shneiderman, 1998), with a special interest in the characteristics and behaviours toward technology of child users (Acuff & Reiher, 1997; Girard & Johnson, 2009; Girard & Johnson 2011; Hanna et al, 1997; Hourcade et al, 2004; Markopoulos et al, 2008; Read et al, 2002; Read, 2008; Scaife et al, 1997). However, there is little in prior research from which recommendations or insights into the best way to evaluate products for children in order to reduce the influence of the evaluation itself on the pedagogical study results. In (Markopoulos et al, 2008) an interesting and useful set of recommendations, tips, and guidelines have been compiled for the evaluation of

interactive products by child-users. This section presents an application of Markopoulos et al's (2008) work for the design of several experimental studies involving children aged seven to eleven from some French and UK primary schools. The field studies, presented in detail in this PhD, were conducted in classroom settings as individual work with technological devices. Some evaluation recommendations were then drawn, from the experience given by all the experimental studies performed during the course of this PhD, extending their work.

- *Factors to consider when designing empirical studies: Theoretical underpinnings.*

The first concern is to define the problem space. Traditionally, this activity begins with the definition of research goals and a set of objectives. It is then followed by a reflection on the choice of technological devices available to the user, the type of study design to select, and a definition of a study procedure. Those considerations should take into account the characteristics of the end users, but the question is: what to choose, how, and when?

In agreement with Markopoulos et al's (2008) recommendations, we consider Piaget's (1952, 1970) work a crucial starting point for the definition of the settings around empirical studies performed on child-users. His work on classifying children by developmental stages includes interesting observations of the way children may behave at the different developmental stages. It also indicates that careful thought should be given to the impact learning can have on the surrounding environment and the child's thought processes.

Our experience of designing and evaluating pedagogical applications for children in this PhD, has involved coupling Piaget's observations with Acuff & Reiher's (1997) stages of child development in interacting with technology. We believe this theoretical foundation enables the design of studies that are more likely to correspond to the children's technological abilities, and cognitive development. Researchers can then choose from a number of alternatives, the most appropriate interaction style (Girard Johnson, 2009; Hourcade et al, 2004; Read et al, 2005), the design and building of the interactive devices (Carroll, 2000; Druin et al, 1999; Druin 2002; Girard & Johnson, 2011), and how to evaluate the adequacy of their products with respect to the user's abilities by selecting relevant evaluation techniques (Dix et al, 2004; Hanna et al, 1997; Markopoulos et al, 2008; Read et al, 2002; Read, 2008; Scaife et al, 1997; Shneiderman 1998).

- *The power of words and consequences of instructions.*

As portrayed by Piaget's theory of development (1952, 1970), children aged seven and older attempt to interpret what is happening around them. This ability is central to many evaluation studies (Hanna et al, 1997; Hourcade et al, 2004), with children entering a complex thought process each time an instruction or evaluation question is given, establishing exactly what is *really* being asked, before constructing an answer or performing a task. For this reason, one has to be careful in how pre-task and post-task briefings, children's interviews, and the instructions given to children for the experimental studies are presented.

For the pre-task briefing in the experimental studies reported in chapters 3 to 7, a choice was made to give the general instructions about the task simultaneously to all participants, in the form of a video projected on the classroom whiteboard. This was to eliminate inconsistencies in instructions and delivery contexts. When working with child-users, it is essential to make sure they understand the tasks to perform, without inferring research goals from the instructions. If not portrayed in a cautious manner, the instructions "you can do X" will be understood by most as either "you shall do X", or "you can do X, but it would be better if you did not", which can alter the study context. A smile, sign of tiredness, or contrariety on the part of the researchers during instructions given to one of the participants has been shown to change the child's entire experience with software, and therefore the study results.

- *Familiarizing yourself with the classroom and the children*

In (Markopoulos et al, 2008), the strong impact of reducing the effect of the children's temperament on the results of the study is highlighted. Several actions are proposed to reduce this effect, including *approach withdrawal*. It is suggested that use of familiar locations and language for the children, will reduce the strangeness of the situation. Additionally it is important to familiarize the children with the people undertaking the study. This facilitates the transfer of the authority from the teacher to the research team for the children to follow instructions more easily, and removes some of the children's shyness towards strangers that can impact the evaluation results.

In our study a month prior to each of the experiments, the researcher directing the study assisted in a mathematics lecture in the classroom, with appropriate consents. The researcher was introduced to the children, along with the topic of the study to be performed, i.e. practice of mathematics tasks. The researcher participated in group-activities with the children, in order for the students to familiarize themselves with her. This step was deemed essential as children aged seven to

eleven can get easily distracted by elements outside of their usual learning environments. By allowing the researcher to be viewed as “another teacher”, and not a stranger, a teacher-child relationship was formed between the children and the researcher. This facilitated children following the researcher’s instructions in a more intuitive way during the time they were alone with her during the study.

- *Naturalistic setting*

In this PhD, the experimental studies were performed using observations of the child participants, in order to minimize the researcher’s influence on the performance of the tasks. Children mostly undertook the tasks individually within the classroom, separate from their peers, in groups of three children. The three participants were placed in the room so as not to see each other work, and in such a manner for the research to naturalistically observe each participant unobtrusively. This setting was found very useful for this PhD’s work, and produced interesting observations of child’s behaviour with a reduced number of interactions between the children and the researcher. It seemed to enable the reduction of the effect of temperament described in (Markopoulos et al, 2008) as *distractibility*, whilst enabling the experiment to be conducted in a timely manner.

Children were brought to the classroom in groups of three. The order of the children to use the software was defined by the teacher. They were placed so as not to see what the other children did. However the researcher was able to take notes on the observations. The children were so involved in the tasks that they forgot the presence of the researcher until the end of the tasks except for occasions where help was needed.

- *Control of the environment*

Another recommendation to reduce the effect of temperament, introduced in (Markopoulos et al, 2008), is *activity level*: making sure that children finishing early do not distract others. Whilst undertaking the study in groups of three children presents advantages, it also can increase the degree of distractibility of children in performing the task, if the study is not planned carefully. When considering the setting up of the study, attention should be given to the number of adults that will be needed to accompany children in and out of the study set-up, and occupy the ones finishing early. Quiet tasks or activities, aiming at occupying children who abandoned the experiment or finished early, should be planned with the children’s teachers prior to the study.

In the study reported in (Markopoulos et al, 2008) children did not take the same time to finish the task. Be respectful enough of the teacher’s planning and timing in class, you need to accompany

all the children back to the classroom together, not individually which would disturb the teacher's flow. However, you do not want to have children wandering around in the room, or left to themselves doing nothing. This is certain to bore them, and try to engage in any activity such as singing, running, fighting, which will end up distracting the others performing the experiment. Make sure you plan this off-time with the teachers with activities they are used to do... drawing activities, puzzles, or have a person read to them quietly in the corner of the room for the youngest ones.

- *Preparing for the experiment: be ready for anything*

In (Markopoulos et al, 2008), the following tips are introduced concerning the evaluation space and to prepare for the study:

“When you plan an evaluation in a space that is not your own, make sure to bring everything you need !” (p.93)

“It is a good idea to visit the test location even before the pilot testing”. (p.93)

This section expands this notion, by giving personal experiences with children using technology, with a non-exhaustive list of things to be aware of, and the possible consequences of not doing so...

In preparing empirical studies, you want to ensure valid, repeatable data from the study set up. However, working with children you will not be able to control as much as with adults. It is really crucial to run a pilot study for many reasons, not least to test all materials needed for the experiment and practical details. This not only includes the technological aspects of the experiment, but most and foremost everything you will need for the experiment to run. You especially need to test things in the final local, if possible, or at least go there and have a look around prior to the experiment to test the feasibility of the study. You have thought of software, the computers, the mice that will be used... think again, that's not all there is to it:

...Use of sound:

In some schools, you have headphones for children to use, however usually of limited numbers and with special plugs that may not work with your material. If you're smart, you might have thought about bringing some headphones from your personal reserves for the experiment... but are you sure they will be usable? Commonly used headphones for portable devices, bright and shiny, really easy to use.. You plug them in and put them on the children's ears... however five minutes in the experiment you will come across a huge problem: they are no good for use by youngsters! The listeners will begin to fall on the ground from five minutes in. As they are too big to be fixed on little children's ears... first, you might try to work it, but you will then have to spare your time between the different children

working at the various stations to help them put them back on, highly distracting for the children in question. What are you going to do tape them on the children's head? Hardly ethical and proper... so you will end up having to use the microphones for hearing, which brings bias in your experimentation as the children will hear what the others will do, and not be able to function isolated from the others anymore.

... plugging in computers and hardware ?

Most of the classrooms in primary schools are not used in being large computer labs, so you might have one or two plugs but not more than that... you therefore would need to take all extension cords and trailing sockets, taping them on the floor. Some researchers advertise the use of SellotapeTM for parcels, which is indeed useful. However, you might also want to plan for less invasive tape, when needed to tape down the alimentation wires of laptop computers on the tables, or sensitive material. This should allow you to leave the school without scorching away the tables and destroying the material.

... timing in the experiment:

Not all classrooms do possess a clock. When the experiment has to follow a specific timing, do not forget to bring a stopwatch for you to regulate the experimental setting. Do not forget to consider the transit time... children will need some time to be taken from their classroom, going to the experimental space, being told the instructions and prepped, performing the experiment, post-session briefing, and then being accompanied back to the classroom. When it is possible, do try to appoint a specific person to the children transit. If not planned carefully, you might end up one group short at the end of the day because it took too much time with the other groups...

...video?

When having written consent from the parents and the children, it is possible and useful to use web cameras embedded in computers to record video data. This enables the participants to "forget" about their being watched during the experiment and to capture more natural settings. However be aware of one humongous problem with using such webcams with children: how short the children truly are when sitting down! Today, you were all set, you have all the technology, the batteries, the wires, plugs, etc.. An adult whose role is to take children in transit, and researches for naturalistic observations... everything should be perfect. First participants come in, sit down, you set up the material, the procedure you rehearsed before the day not to loose time... and oops... you only half of the child's head!!! Indeed in primary schools the settings are low in the chair-table proportion. This does not enable you to see the child's head when working on it... One cheap solution: you need to be

not shy of robbing your apartments for all cushions to set the children safely and efficiently so as to see what you will be looking for in the analysis.

- *Rewards*

When working with users, it is recommended to thank all participants as soon as possible after the evaluation exercise is completed. In (Markopoulos et al, 2008), it is suggested to offer a small token, such as a “certificate” to the participants as a reward. Children are known not to really need rewards to perform the studies, as they are usually looking forward to doing tasks for adults. However, the fact to give a “certificate”, as a souvenir of its participation to the project, makes it possible for the children to share what their experience was with their parents. In one of the experiment, before leaving, the researcher thanked again all child-participants and the teacher, and gave to the class a number of personalized gold stars, bearing the logo of *Multiplotest*, as organized with the teacher beforehand. In this school, stars are earned by children upon excellent scholastic results, good behaviour, helping others etc... and they grant “special treats” for the entire class when enough are earned. If possible it is a good idea to work something out with the teacher as part of the class practices. This practice indeed reinforced the positive experience of children in doing the experiment, and facilitated the reused of the participants for other projects ongoing.

In this section, a set of evaluation recommendations were introduced, illustrated by personal experiences in the design and realization of empirical studies for children aged seven to eleven years olds working on mathematical pedagogical applications. The work of Markopoulos et al (2008), among others, was extremely valuable for the initial design choices to be made in the process of designing experimental studies for pedagogical applications. Practical examples were given before as to the possible improvements of empirical study design in an effort to reduce usability issues that can hinder the investigation of pedagogical research questions. The set of recommendations listed here, while far from exhaustive, aims at helping future researchers in child educational technology in their experiment planning. However, there is still room for investigations as to the impact of technology and technological settings on final users, and how this can affect the realization of the studies and children’s experience in terms of learning.

8.2.6 CHILDREN'S USE OF THE LEARNER MODEL

Finally, the main contribution of the thesis is the empirical claim that affective OLM tutors can be used with children aged seven to eleven years for better mathematical learning. The research questions summarized in the first part of the chapter have been addressed by investigating existing theories, practical usability studies, and empirical testing, in consideration of the following aspects: how to represent an learner model that is suitable for children to use, how children view the benefits of accessing such information, what mechanism can help them use it better, and to what extent they can make use of said model.

As discussed in the literature review (section 1.4.3) and in chapter 4, learner models have been designed using all sorts of visual representations, and differ in the information considered to be part of the OLM application. Several questions are considered when producing OLM pedagogical software: What information from the learner model can be inspected, and modified, and by whom? What does a learner model component look like, and how can one represent a set of learner model components, in order to give the user an understanding of the model in its whole? In order to represent a LM, two design choices have to be made: the selection of a number of LM views, and a number of LM visualization tool to be used (cf. chapter 4 for a definition of the two concepts). The use of multiple representations of the LM in a single application has proven to be useful for adult and adolescent students to access and understand the model in the pedagogical software (Zapata-Rivera & Greer, 2004; Mabbott & Bull, 2004). When considering child-users, a few visualization tools have been used from simple skill-meters and boxes (Bull & McKay, 2004) to Bayesian graphs (Zapata-Rivera, 2003), usually under a 'concept-map' or 'related concepts' view.

In chapter 4, two usability studies were performed in order to answer the following research question: *what is the best representation of a learner model for children aged seven to eleven years to understand and interact with?* The studies, involving 100 children aged seven to eleven years, showed a difference in children's preferences, and understanding, of the visualization tools and LM views.

In study 1, children aged seven to eleven seem to prefer 'skill-meters' and 'boxes' tools, with 'Bayesian graphs' being the least favoured tool, and seem to held a preference for 'lecture' view, favouring the least the 'pre-requisites' view. There seem to be some difference in terms of understanding of the underlying LM concepts according to the age of the participants: The younger children group (seven to nine) seem less able, or willing, to explain a learner model when represented in text. The results of the first study show that children are most likely to understand better the LM tools as they age: indeed, as they progress in school, the way the pedagogical information visualization

is portrayed evolves, from pictorial, to textual, to graphical representations, with the introduction of graphs and tables in history/geography around eight years of age in French schools.

In study 2, no analysis of the understanding of the LM according to the age of the participants was performed. Indeed, as the future OLM applications to be designed were to be used by children aged seven to eleven, the study investigated the views best recognized by all, and not by only one of the age groups. Some difference appeared in the general understanding of LM represented using the different LM views: two of the four tested ('lectures', 'related concept', 'concept map', and 'pre-requisites') seemed better understood by children aged seven to eleven: the 'lectures' and 'related concept' views. Under such views, there is a consistency of higher level of understanding of the 'skill-meter' and 'boxes' visualization tools, when compared to the other tools investigated ('text', 'table', '2D-graph'). From the results of these studies, the learner model of the *DividingQuest* and the different versions of *Multipliotest* were created using 'skill-meters' and 'boxes' as visualization tools, under both a 'lectures' and 'related concepts' view. Children seemed to favour, be more motivated to use, and understand better, the learner model when its representation in the form of a 'skill-meter' or 'boxes' was displayed using an embodied pedagogical character, portraying emotional feedback. The fact to embody the representation of the LM component, and to give the learner the opportunity to argue with the character about it's the actual representation of what s/he knows, seemed to motivate the child in inspecting the model content, and be more willing to engage in a negotiation process to built a joint-representation of the model (system and child's view of the achievements).

The empirical evaluations of OLM applications in the later chapters showed a good understanding and use of the LM under the views and tools used. However, the personalization of the LM to a child's preferences was not permitted, and the use of multiple vs. single representations of the LM was not investigated. A further analysis of the various uses of embodied pedagogical agents in the design of LM visual representations, and the need for the use of multiple views and representation tools for young users to manipulate OLM applications might therefore be of great research value.

Children aged seven to eleven seem to be willing to use LM information to inform their actions on software, in particular when confronted to Negotiated-Learner Model applications, that proposes a representation of the model easily understood. Results from chapter 5's study show that the presence or absence of learner model information seems to have an impact on how enjoyable the children found the mathematical software, as well as children's learning: children using the Open-Learner Model version seemed more satisfied with the application, experiencing more enjoyment out of the game, and visualizing more how it could be used to learn maths than the ones working on the Closed-Learner Model version. They also performed better on the OLM condition than on CLM software. Children wished to be extend the information they were able to access, and to gain more control over the information presentation, and edition. They wished, in a way similar to Bull et al's

(2005) and Zapata-Riviera's (2004) results on the use of OLM application with children and teenagers, to be able to access representations of the concepts they understand, the ones they have not fully grasped yet, and misconceptions or a listing of errors they currently make.

The use of affective feedback given by Embodied-Pedagogical Agents to inform on the learner model content, and help interact with the model, seemed to help children take a greater interest in the LM, and how it could help them learn. In chapter 6, using an EPA with evolutionary affective state according to the users software interaction to interact with the content of the learner model is more beneficial than using a mode "traditional EPA (which expression does not vary from a fixed happy smile), or an interface where the interaction is only led by verbal feedback and the use of traditional user interface components and LM representation tools in terms of children's motivation to use software, the number of times the model was inspected, and the definition of a more 'accurate' learner model.

Children want to access their learner model components, and such access, as well as the understanding of the learner model content, is facilitated by affective EPAs. These results led to the last research question investigated: how are children willing to access and edit learner model information, and what is the impact on learning of such inspection/edition? For this study, three types of OLM applications were compared, with the model separated into the system's and children's assessment of their skills: an *inspectable* (system's view only), an *editable* (child's view editable, no inspection of the system's view), and at last, a *negotiated* learner model (where a negotiation process is elicited in order to construct the model using the child's and the system's beliefs about the child's current knowledge).

Children seemed to all find some interest in using the learner model, to different extent, and for different periods of time:

- Children working on the *inspectable learner model* found the display information about the system's assessment useful, especially to guide them in which activity to perform next in relation to the system's feedback.
- Children working on the *editable learner model*, with no representation of the system's assessment, seemed to enjoy working on the interface the first time, but progressively consider the application less useful, or even "boring" in the following sessions. Some of them went back to the learner model display to update the changes in their knowledge that could happen by performing the activities, but most of the children stopped making modifications after a few activities.
- Children working on the *negotiated learner model* seemed to use the information on the learner model provided to inform their next action. After exploring the software for some time, they went directly to the learner model presentation. They then seemed

to look for the concepts where a disagreement could be graphically seen between the child's and system's view of the learner model.

The use of affective feedback given by Embodied-Pedagogical Agents, that serve as interaction medium with the learner model, has proven to work best for children when embedded in a Negotiated-Learner Model tutor, where the model is jointly-constructed. The use of two different representations of the learner model content - one representing the children's self-beliefs, and the other the system's assessment of knowledge acquisitions – has proven to lead children to be more involved in the representation of what they know by visually comparing their views of how much a specific concept is grasped to the system's assessment, and engaging in a negotiation process when a disagreement was found.

Children working on the s-LM and NLM interfaces were sensitive to the affective feedback given by the EPA representing the system's assessment during an activity or in the representation given of the model. They reacted strongly, emotionally, to the under-estimation of their abilities by the system, and this influenced their next action in software: showing the EPA representing the system's view that they were right about their assessment. The characters became alive, a representation of sort some of expert judge, or the teacher's voice, to whom they “had to prove” that their assessment was correct.

There seems to be a developmental difference in the way children edited the learner model. It seems that while the seven to nine years were as keen as the nine to eleven years to go and inspect the model content, as an inspiration for future actions decision, they revealed to participate less in the join-construction of the learner model. Indeed, when confronted with a disagreement between their beliefs and the system's assessment, the majority of youngest children preferred to agree with the system's decision, rather than engaging in a negotiation process to prove their view. This behaviour might be explained by the way children conceive what the system's view represent. In the post-session interviews, it appeared that *Multiplio*, the system's representation of the LM through an EPA, was considered *alive* by children: a character capable of rational thought and expressing feelings. For most of the children involved in the study, *Multiplio* then represented what the teacher thinks of their achievements. Children aged seven to nine are not used to argue learning assessments provided by teachers, and therefore might have preferred a less ‘confrontational’ approach of disagreeing with the system. Indeed, when the system's representation of one skill was deemed overly under-estimated, they expressed vocally their astonishment or even affront, and made a point in choosing the activity to “prove the character wrong”. The eldest children, when confronted to this situation, were more keen on explaining their view and, when an agreement couldn't be made, go back to the activity to justify their assessment.

8.3 FUTURE WORK

8.3.1 APPLICATIONS AND FEASIBLE RESEARCH DIRECTIONS WITH THE CURRENT SYSTEMS

- Varying the emotional response of the EPAs

As described earlier, there is a developmental difference in the way children understand and represent emotional states. In this PhD, a design choice was made to create a model of affective feedback (A.M.E.R.) representing only unique emotions. The A.M.E.R. model was only tested against a lack of feedback in chapter 5, but not against other models of emotion. There would be a need to validate the definition of the three strategies of affective feedback in a larger-scale experiment than the pilot study detailed in chapter 3. The model should then be tested against other theories of emotion, such as the strategy used in the Wayang Tutor (Woolf et al, 2010), and tested in terms of learning gain, and use of the OLM application. Another research issue would be the investigation into how the representation of multiple emotions, or a succession of unique emotions in the model might impact the way children react to the feedback, when the younger children are only aware of unique emotions, and the elder children of multiple emotions.

- Testing the applicability of the affective model on different children age groups and nationality, or problem domain.

The A.M.E.R. model was built and aimed at French and English children aged seven to eleven years, working on drill-and-practice mathematical applications. As discussed in the literature review, and portrayed by the study results, there is a developmental difference in children's understanding of emotions, and recognition of the moment of elicitation. While the selection of emotional responses was defined with the pedagogical strategies and emotional development stages of such age and country of schooling groups, the extendibility to other nationalities and older children should be investigated. Similarly, the extendibility of the model to other pedagogical concepts should be investigated. The instructions defined in the model clearly correspond to a drill-and-practice application, but may not be necessarily linked to mathematical concepts only.

- Representing a learner model for children's use: which representation to offer users from different age-groups?

The study in chapter 4, and the empirical observations of children's use of the LM in the DividingQuest and Multipliotest confirmed a good understanding of the learner model when represented using a "skill-meter" or "boxes", under a "lectures" or "related-concepts" view. Chapter 7 shows an increase in the level of child's motivation to inspect/modify the model, and actual use of the LM information when this information is portrayed through the use of an embedded pedagogical agent, rather than a "traditional" LM representation, such as in chapter 5. While in this thesis's empirical evaluations, the choice of one representation of the learner model for all types of users was founded, it may not be the case of other applications: the first study showed that the understanding of the representation tools varied according to the age of the participants. Is it therefore beneficial for learning to restrict the visualization of the model to a form known by all, or would the choice to personalize the way in which the information is presented to the user help gain a greater level of interest in the model content and manipulation? Similarly, children's use of a single, or multiple views and tools to inspect and modify the model has not been part of this PhD's work, but its impact on children's performance, and use of the model, could be investigated further.

8.3.2 LONG TERM RESEARCH DIRECTIONS

- Improving the maintenance of a jointly constructed LM.

The results of the last experiment, in chapter 7, show children's willingness to work in cooperation with the system, to build a more accurate view of his/her competency level, in the form of a jointly-constructed learner model. The current learner model component is quite simple in the definition of the different concepts relating to learning, and how learning evolves. The use of more sophisticated underlying learner model could help in reducing errors in system's assessment of children's skills. The negotiation process involved here did not use natural language processing as some of the children involved were deemed having too low a literacy skill to enable this process. Different ways to negotiate with the model in order to produce a jointly constructed model, from persuasion, to giving more evidence, should be investigated for children this age group.

- Investigating teachers' use of such OLMs.

When deploying the Multipliotest and DividingQuest applications into schools, the teachers participating in the experiments gave some comments on the possible use of the tutors as assessment tools in addition to the current use by child-users. The results of the experiment in chapter 7 show children's willingness to inform the system about their skill acquisition, as well as their evolution of emotional states. Such information could therefore be accessible, and to some extent editable, by other actors of the learning process, such as the child's teacher. The teacher's interface, currently enabling only one view of the child's achievements and learner model use, should be extended to include all aspects of the learner model edition and joint-construction in the different interfaces, and provide teachers with interfaces that facilitate children's assessment in terms of skills acquired and misconceptions found by software use.

- Extending the use of the proposed OLM tutors to other learning domains and learning styles.

Finally, this work has been performed on drill-and-practice mathematical software. The extendibility of the results on other types of concepts to be learned, and underlying learning styles should be investigated further. Indeed, while the DividingQuest and Multipliotest software were created in participatory-design with teachers and children of the end-user's age groups and nationality, the generalization of the results in terms of willingness to use OLM applications and the understanding of affective components, might be limited to French and English children aged seven to eleven years. It might be interesting to investigate to which extent the results found in this PhD's work are limited to the building of mathematical drill-and-practice applications, for young children, and how replicable the design, evaluation, and use of interactive products hereby presented, are.

References

Web references:

- DFEE, 1999 : Mathematics, The National Curriculum for England. Available from <http://curriculum.qcda.gov.uk/key-stages-1-and-2/subjects/mathematics/keystage2/> Accessed on January 14th 2009.
- PNS, 2007: Primary Numeracy Strategy: Overcoming barriers in mathematics – helping children move from level 3 to level 4, Ref: 00695-2007PCK-EN, available from www.standards.dcsf.gov.uk, accessed on July 17th 2011.

Conference articles, Journal articles, and Book references:

- 8fish, 2008. *Making faces: drawing expressions from comics and cartoons*. Cincinnati:f+W media.
- Acuff, D. S., and Reiher, R. H., 1997. *Kids' Stuff: Toys and the changing worlds of American childhood*. Harvard: Harvard University Press
- Ahmad, N. & Bull, S. 2008. Do Students Trust their Open Learner Models?, In W. Nejdl, J. Kay, P. Pu & E. Herder (eds), *Adaptive Hypermedia and Adaptive Web- Based Systems*, Berlin Heidelberg: Springer-Verlag, pp.255-258.
- Ainley, M., Being and Felling interested: transient state, mood, and disposition, In P.A. Schutz & P. Pekrun (Eds.), *Emotion in education*. San Diego:Academic Press, pp 147 – 164.
- Alessi, S. M., & Trollip, S. R., 2001. *Multimedia for Learning. Methods and development*. Massachussetts: Allyn&Bacon.
- Aleven, V., Popescu, O., & Koedinger, K.R. 2001. Towards tutorial dialog to support self-explanation: Adding natural language understanding to a cognitive tutor. In : *Proceedings of the 10th conference on Artificial Intelligence in Education, San Antonio, Texas* , pp.246-255
- Alsmeyer, M., Luckin, R., and Good, J. 2007. Getting Under the Skin of Learners: Tools for Evaluating Emotional Experience. In *Proceeding of the 2007 Conference on Artificial intelligence in Education: Building Technology Rich Learning Contexts that Work* R. Luckin, K. R. Koedinger, and J. Greer, Eds. *Frontiers in Artificial Intelligence and Applications*, vol. 158., Amsterdam, The Netherlands: IOS Press, pp.153-160.
- Ames, C. 1992. Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261-271.
- André, E. , T. Rist , J. Müller, Integrating reactive and scripted behaviors in a life-like presentation agent, In: *Proceedings of the second international conference on Autonomous agents*, May 10-13, 1998, Minneapolis, Minnesota, United States, p.261-268
- Arroyo, I., et al, 2010. Improving Math Learning through Intelligent Tutoring and Basic Skills Training, In V. Aleven, J. Kay and J. Mostow (Eds.) *Intelligent Tutoring Systems: 10th International Conference, ITS 2010. LNCS 6094*. Berlin Heidelberg: Springer-Verlag. pp. 423-443
- Aspinwall, L. G., & Taylor, S. E. (1997). A stitch in time: Self-regulation and proactive coping. *Psychological Bulletin*, 121, 417-436.
- Baecker, R., Small, I., and Mander, R. 1991. Bringing icons to life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through*

- Technology* (New Orleans, Louisiana, United States, April 27 - May 02, 1991). S. P. Robertson, G. M. Olson, and J. S. Olson, Eds. CHI '91. ACM, New York, NY, 1-6.
- Baker, R., Corbett, A.T. and Koedinger, K. (2004) 'Detecting student misuse of intelligent tutoring systems', *Proceedings of the Seventh International Conference on Intelligent Tutoring Systems*, pp. 531-540.
- Barnard, Y., and Sandberg, J. 1996) Self-explanations, Do We Get them from Our Students? *Proceedings of European Conference on Artificial Intelligence in Education*. Lisbon , Spain.115-121
- Barnard, Y.F., Erkens, G. & Sandberg, J.A.C. 1990. Interaction in intelligent tutoring systems. *Journal of Structural learning*, 10(3), 197-213.
- Barrett, L.F., and J.A.Russel, 1998. Independence and bipolarity in the structure of current affect, *Journal of Personality and Social Psychology*, 74(4), 967-984.
- Bates, J. (1994). "The role of Emotion in Believable Agents". *Communications of the ACM* 37(7): pp 122 - 125.
- Batson, C., Turk, C., Shaw, L., & Klein, T., 1995. Information function of empathic emotion: Learning that we value the other's welfare. *Journal of Personality and Social Psychology*, 68, 300-313
- Baylor, A. 2005. The impact of Pedagogical Agents on affective outcomes. In *Proceedings of the Workshop on Affective Interactions: Computers in the Affective Loop*. 10th International conference on Intelligent User Interfaces, 2005, San Diego, CA.
- Baylor, A., & Kim, Y. 2005. Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education*, 15(1), 95-115.
- Bekker, M., J. Beusmans, D. Keyson and P. Lloyd, 2002. KidReporter: a method for engaging children in making a newspaper to gather user requirements, *Proceedings of Interaction Design and Children 2002* (Eindhoven, NL, 2002) Eindhoven:Shaker.
- Berkovitz, J. 1994. Graphical interfaces for young children in a software-based mathematics curriculum. In *Proceedings of the ACM Conference on Human Factors in Computing Systems: Celebrating Interdependence*, CHI '94, Boston, MA, Apr. 24-28. ACM Press, New York, NY, 247-248.
- Bickmore, T., Picard, R.W. 2004. Establishing and Maintaining Long-Term Human-Computer Relationships. *Transactions on Human-Computer Interaction* 12, 293-327 2004.
- Bloom, L. 1998. Language acquisition in its developmental context. In: W. D. K. Siegler (Ed.), *Handbook of child psychology: Vol2. Cognition, perception, and language*. New York: Wiley. pp. 309-370
- Boekaerts, M., 2007. Understanding Student's affective processes in the classroom. In P.A. Schutz & P. Pekrun (Eds.), *Emotion in education*. San Diego:Academic Press, pp. 37-56
- Bosenberg, A., Thomas, J. et al. 2003. Validation of a six-graded faces scale for evaluation of postoperative pain in children. *Paediatric Anaesthesia*. 13, 8 (Oct. 2003), 708-713.
- Boud, D., Keogh, R., & Walker, D. (1996). What is reflection in learning. In D. Boud, R. Keogh & D. Walker (Eds.), *Reflection: turning experience into learning* (pp. 7-17). London: Kogan Page.
- Bradley, J., 1995. A model for evaluating student learning in academically based service. In Troppe, M. (Ed.), *Connecting cognition and action: Evaluation of student performance in service learning courses*. Providence., RI. pp13-26
- Brave, S., Nass, C., & Hutchinson, K. 2005. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, 62(2), 161-178.

- Broekens, J. and W.P. Brinkman, 2009. AffectButton: Towards a Standard for Dynamic: Affective User Feedback. In ACII 2009, IEEE.
- Brown, J. S., & Burton, R. R. 1978. Diagnostic Models for Procedural Bugs in Basis Mathematical Skills. *Cognitive Science*, 2, 155-191
- Brown, J.S., Burton, R.R. & Bell, A.G. 1975. SOPHIE: A step toward creating a reactive learning environment. *International Journal of Man-Machine Studies*, 7, 675- 696.
- Bruce, T., 2001. *Learning through Play: babies Toddlers and the. Foundation Years*. London: Hodder and Stoughton
- Bruner, J., 1966. *Studies in cognitive growth : A collaboration at the Center for Cognitive Studies*. New York: Wiley & Sons..
- Brusilovsky, P. 1994. Student model centered architecture for intelligent learning environments. In Proceedings of Fourth international conference on User Modeling, User Modeling Inc, 31-36.
- Brusilovsky, P. 1995. Intelligent learning environments for programming: The case for integration and adaptation. Proceedings of the 7th World Conference on Artificial Intelligence in Education.
- Brusilovsky, P. & Sosnovsky, S. (2005). Engaging Students to Work with Self-Assessment Questions: A study of two approaches, Proceedings of 10th Annual Conference on Innovation and Technology in Computer Science Education, ACM Press, 251-255.
- Bull, S. & Kay, J. 2005. A Framework for Designing and Analysing Open Learner Modelling, Proceedings of Workshop on Learner Modelling for Reflection, International Conference on Artificial Intelligence in Education, 81-90.
- Bull, S., & Kay, J. 2007. Student Models that Invite the Learner In: The SMIL© Open Learner Modelling Framework. *International Journal of Artificial Intelligence in Education*, 17(2), 89-120.
- Bull, S. & Kay, J. 2010. Open Learner Models, In: R. Nkambou, J. Bordeau & R. Miziguchi (eds), *Advances in Intelligent Tutoring Systems*, Springer, pp.318-338.
- Bull, S., & Mabbott, A., 2006. *20000 Inspections of a Domain-Independent Open Learner Model*. ITS 2006, LNCS 4053, pp. 422 – 432, 2006.
- Bull, S. & McEvoy, A.T. (2003). An Intelligent Learning Environment with an Open Learner Model for the Desktop PC and Pocket PC, in U. Hoppe, F. Verdejo & J. kay (eds), *Artificial Intelligence in Education*, IOS Press, Amsterdam, 389-391
- Bull, S. & McKay, M. 2004. An Open Learner Model for Children and Teachers: Inspecting Knowledge Level of Individuals and Peers, In: J.C. Lester, R.M. Vicari & F. Paraguacu (eds), *Intelligent Tutoring Systems: 7th International Conference*, Springer-Verlag, Berlin Heidelberg, 646-655.
- Bull, S., & Pain, H. 1995. 'Did I say what I think I said, and do you agree with me?': *Inspecting and Questioning the Student Model*. Paper presented at the World Conference on Artificial Intelligence in Education, AACE.
- Bull, S., Cooke, N. & Mabbott, A. (2007). Visual Attention in Open Learner Model Presentations: An Eye-Tracking Investigation, in C. Conati, K. McCoy & G. Paliouras (eds), *User Modeling 2007: 11th International Conference*, Springer-Verlag, Berlin Heidelberg, 187-196
- Bull, S., Gardner, P., Ahmad, N., Ting, J. & Clarke, B. 2009. Use and Trust of Simple Independent Open Learner Models to Support Learning Within and Across Courses, in G-J. Houben, G. McCalla, F. Pianesi & M. Zancanari (eds), *User Modeling, Adaptation and Personalization*, Springer-Verlag, Berlin Heidelberg, 42-53.
- Bull, S., Mangat, M., Mabbott, A., Abu Issa, A. S., & Marsh, J. 2005. Reactions to Inspectable Learner Models: Seven-Year Olds to University Students. Paper presented

- at the 12th International Conference on Artificial Intelligence in Education (AIED 05), Amsterdam.
- Burleson, W., Picard, R.W., 2007. Gender-Specific Approaches to Developing Emotionally Intelligent Learning Companions. *IEEE Intelligent Systems* 22, 62–69
- Butler, R. 1992. What young people want to know when: Effects of mastery and ability goals on interest in different kinds of social comparison. *Journal of Personality and Social Psychology*, 62, 934-943.
- Capell, P. and Dannenberg, R.B. 1993. Instructional design and intelligent tutoring: Theory and the precision of design. *International Journal of Artificial Intelligence in Education*, 4,95-121
- Carr, B., Goldstein, I. 1977. *Overlays: a Theory of Modelling for Computer Aided Instruction*. AI Memo 406, MIT, Cambridge, USA.
- Carroll, J. 2000. *Making Use: Scenario-Based Design of Human-Computer Interactions*. MIT Press.
- Carver, C., & Scheier, M. 1990. Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, 97, 19-35.
- Carver, C.S., Sutton, S.K., & Scheier, M.F., 2000. Action, emotion, and personality: Emerging conceptual integration. *Personal and Social Psychology Bulletin*, 26, 741-751.
- Chan, T. 1996. Learning Companion Systems, Social Learning Systems, and the Global Social Learning Club. *International Journal of Artificial Intelligence in Education*. 7(2), 125-159.
- Chan, T. W., & Baskin, A. B. 1990. Learning Companion Systems. In Frasson, C., and Gauthier, G. (eds.) *Intelligent Tutoring Systems: At the crossroads of AI and Education*. Ablex Pub. Norwood, NJ, 6-33.
- Chiasson, S., & Gutwin, C., 2005a. *Design for Children's Technology*. HCI-TR-2005- 02.
- Chiasson, S., & Gutwin, C. 2005b. *Testing the media equation with children*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.
- Corbett, A.T. and Anderson, J.R. 1995. Knowledge ‘racing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interactions* 4, 253-278
- Corbett, A. T., & Bhatnagar, A. 1997. *Student Modeling in the ACT Programming Tutor: Adjusting a Procedural Learning Model with Declarative Knowledge*. Paper presented at the User Modeling: Proceedings of 6th International Conference.
- Corbett, AT., Anderson, J.R. and Patterson, E.G. 1990. Student modeling and tutoring flexibility in the Lisp Intelligent Tutoring System. In C. Frasson and G. Gauthier (Eds.) *Intelligent tutoring systems: At the crossroads of artificial intelligence and education*. Norwood, NJ; Ablex Publishing.
- Craig, S., Gholson, B., & Driscoll., D. 2002. Animated pedagogical agents in multimedia educational environment: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology*, 94(2).
- D'Mello, S. K., S.D. Craig, A.M. Witherspoon, B. McDaniel, and A. C. Graesser 2008): “Automatic detection of learner's affect from conversational cues”. *User Model User-Adaptation and Interaction* 18(1-2), 2008, 45-80
- Damasio, A., 1995. *L'erreur de Descartes. La raison des émotions*. Paris : Édts Odile Jacob (Sciences) .
- Darwin, C. R., 1872. *The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray. 6th edition; with additions and corrections. Eleventh thousand.

- Darwin, C., 1998. *The expression of the emotions in man and animals*. 3rd edition. New York: Oxford University Press.
- Denham, S.A. and Auerbach, S. 1995. Mother-child dialogue about emotions and preschoolers' emotional competence. *Genetic, Social, and General Psychology Monographs*. 121, 3 (Aug. 1995), 311-337.
- Dillenbourg, P. 1996. Some Technical Implications of Distributed Cognition on the Design on Interactive Learning Environments. *International Journal of Artificial Intelligence in Education*. 7 (2), 161-179.
- Dimitrova, V. 2003. StyLE-OLM: Interactive Open Learner Modelling. *International Journal of Artificial Intelligence in Education*, 13(1), 35-78.
- Dimitrova, V., Self, J., & Brna, P. 2001. Applying Interactive Open Learner Models to Learning Technical Terminology. *Proceedings of the 8th UM 2001*, Springer, pp. 148-157.
- Dimitrova, V., McCalla, G., & Bull, S., "Open Learner Models: Research Questions (Preface of the Special Issue, Part 1.)" *International Journal of Artificial Intelligence in Education*, 17(2), 2007, 83-87.
- Dix, A., Finlay, J., Abowd, G., & Beale, R., 2004. *Human Computer Interaction*, 3rd edition: Prentice Hall
- Donaldson, M., 1978. *Children's Minds*, New York: Fontana.
- Draper, S. W. 1999. Analysing fun as a candidate software requirement. *Personal Technology*, 3, 117-122.
- Druin, A. 1999., Cooperative inquiry: developing new technologies for children with children, In: *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, p.592-599, May 15-20, 1999, Pittsburgh, Pennsylvania, United States
- Druin, A., & Inkpen, K. 2001. *When are Personal Technologies for Children?* Paper presented at the Personal and Ubiquitous Computing.
- Druin, A., Stewart, J., Proft, D., Bederson, B., Hollan, J. (1997). KidPad: A design collaboration between children, technologists, and educators. *Proceedings of CHI'97*, Atlanta, GA.
- Druin, A., Weeks, A., Massey, S., & Bederson, B. B. (2007). Children's interests and concerns when using the International Children's Digital Library: A four country case study. In *Proceedings of Joint Conference on Digital Libraries (JCDL'2007)* Vancouver, British Columbia, Canada.167-176.
- Dweck, C. & Leggett, E. 1988. A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.
- Economides, A., & Elissaver, G. 2000. *Evaluation Factors of Educational Software*. Paper presented at the IEEE.
- Ekman, P. 1972. *Universal and cultural differences in facial expressions of emotion*. Paper presented at the Nebraska symposium on motivation.
- Ekman, P., 2007, *Emotions revealed: recognizing faces and feelings to improve communication and emotional life*, 2nd edition, Holt Paperbacks, New York, 290p.
- Ekman, P., Friesen, W., 1978. *Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P. & Friesen, W.V., & Ellsworth, P. 1972. *Emotion in the human face : guidelines for research and an intégration of findings*, New York : Pergamon Press, 1972.
- Ekman, P., Friesen, W.V., & Tomkins, S.S. 1976. *The facial affect scoring technique (FAST)*, Palo Alto, Calif : Consulting Psychologists Press, 1976.
- Eibl-Eibesfeldt, I., 1971. *Love and hate : on the natural history of behavior patters*. New York: Holt, Rinchart & Winston.

- Eibl-Eibesfeldt, I., 1973. The expressive behavior of the deaf-and-blind born. In: M. von Granach & L. Vine (Eds.) *Social communication and movement*. New York: Academic Press.
- Estrada, C.A., Isen, A.M., Young, M.J., 1994: Positive affect influences creative problem solving and reported source of practice satisfaction in physicians. *Motivation and Emotion* 18, 285-299 1994.
- Felder, R. M. and Silverman, L. K.: Learning and Teaching Styles in Engineering Education. *Engineering Education*, 78(7) (1998) 674-681.
- Fisher, R. 1995. *Teaching children to learn* London: Blackwell/Simon & Schuster /Stanley Thornes
- Fontijn, W. and J. Hoonhout. Functional Fun with Tangible User Interfaces. In Proc of DIGITEL2007. IEEE Computer Society, Washington, DC, 2007.
- Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology*, 2, 300–319.
- Frijda, N. J. 1986. *The Emotions*. Cambridge: Cambridge University Press.
- Gallimore, R. & Tharp, R. 1990. Teaching mind in society. In L. Moll (Ed.). *Vygotsky and education: Instructional implications and social applications of socio-historical psychology*. New York: Cambridge University Press. pp. 175-205
- Gilutz, S., and Nielsen, 2002. *Usability of websites for children: 70 design guidelines*. Preemont, CA: Nielsen Norman Group.
- Goleman, D. 1995. *Emotional Intelligence*. New York :Bantam Books.
- Gordon, J. (1996). "Tracks for Learning: Metacognition and Learning Technologies." *Australian Journal of Educational Technology* 12(1): 46-55.
- Graham, S., Hudley, C., & Williams, E. 1992. Attributional and emotional determinants in aggression among Afro-American and Latino young adolescents. *Developmental Psychology*, 28(4), 731-740.
- Graham, S., Weiner, B. (1996). Theories and principles of motivation. In *Handbook of Educational Psychology*. Berliner & Calfee (Eds.). New York, NY: Simon & Schuster Macmillan.
- Greig, A. and A. Taylor, *Doing Research with Children*. 1999, Sage. London: 1999,
- Grigoriadou, M., Tsaganou, G., & Cavoura, T. 2003. *Dialogue-Based Reflective System for Historical Text Comprehension*. Paper presented at the International Conference on Artificial Intelligence in Education.
- Guha, M.L., et al. Mixing Ideas: A New Technique for Working with Young Children as Design Partners In Proc of IDC 2004. ACM Pres, Maryland, USA, 2004
- Gulz, A. 2004, Benefits of virtual characters in computer based learning environments: claims and evidence. *International Journal of Artificial Intelligence in Education*, 14(3), 313-334.
- Haake, M. and Gulz, A. 2009 A Look at the Roles of Look & Roles in Embodied Pedagogical Agents - A User Preference Perspective. *International Journal of Artificial Intelligence in Education*, 19(1), 39-71.
- Hammond, J., 2001. *Scaffolding: Teaching and learning in language and literacy education*. Newtown, Australia: PETA.
- Hass, H. 1970. *The human animal: the mystery of man's behavior*. New York: Putnam
- Hall, L., Jones, S., & Hall, M. 2008. *Investigating User Engagement with Synthetic Characters*. In Proceedings of International Conference on Advanced Learning Technologies 2008. Eighth IEEE International Conference on.
- Hanna, L., Ridsen, K., & Alexander, K. 1997. Guidelines for usability testing with children. *Interactions*, 4(5), 9-14.

- Hanna, L., Ridsen, K., Czerwinski, M., & Alexander, K. J., 1998. The role of usability research in designing children's computer products. In *The design of children's technology* Morgan Kaufmann Publishers Inc. pp. 3-26
- Harris, P. L., 1989. *Children and Emotion*. Oxford: Blackwell Publishing.
- Harris, P. L. 1997. Between strange situations and false beliefs: working models and theories of mind. In J. H. D. v. d. B. W.Koops (Ed.), *Early mother-child interaction and attachment: Old and new approaches*. Amsterdam: Royal Netherlands Academy of Arts and Science. pp. 187-199
- Harter, S. (1983). Developmental perspectives on the self-system. In P. H. Mussen (Ed.), *Handbook of Child Psychology*, (Volume IV, 4th edition, pp. 275-385). New York: Wiley
- Harter, S., and Buddin, B., 1987. Children's understanding of the simultaneity of two emotions: a five-stage developmental acquisition sequence, *Developmental psychology*, 23, 388-99.
- Harter, S., & Whitesell, N. R. 1989. Developmental changes in children's understanding of single, multiple, and blended emotion concepts. In C. S. P. Harris (Ed.), *Children's understanding of emotions*. Cambridge: Cambridge University Press. pp. 81-116
- Hatcher, J.A. & Bringle, R. G. 1997. Reflection: Bridging the gap between service and learning. *College Teaching*, 45(4), 153-158.
- Heckhausen, H., 1991. *Motivation and action*. New York: Springer.
- Hedberg, J. G., Harper, B., Brown, C., and Corderoy, R. M. 1994. « Exploring user interfaces to improve learning outcomes ». In Proceedings of the IFIP Tc3/Wg3.2 Working Conference on the Seign, Implementation and Evaluation of interactive Multimedia in University Settings: Designing For Change in Teaching and Learning (July 06 - 08, 1994. K. Beattie, C. McNaught, and S. Wills, Eds. IFIP Transactions, vol. A-59. Elsevier Science, New York, NY, 15-29.
- Henderson, H.A., & Fox, N.A., 2007. Considerations in studying emotions in infants and children. In Coan, J.A. & Allen, J.J.B., *Handbook of emotion elicitation and assessment*, Oxford: Oxford university press, pp349-361.
- Heraz, A., Frasson, C., 2008. "Detecting the Affective model of interplay between emotions and learning by measuring learner's brainwaves." WECITS'2008 (pp55-62. ITS 2008. Montreal, Canada, June 2008.
- Holbrook, M. B. 1985. Emotion in the consumption experience: Toward a new model of the human consumer. In H. a. W. E. Peterson (Ed.), *The role of affect in consumer behavior: Emerging theories and applications*. Lexington: MA:Heath.
- Hopkins, C., Gifford, S., & Pepperell, S. 1996. *Mathematics in the Primary School*: David Fulton Publishers.
- Hourcade, J. P., Crowther, M., and Hunt, L. 2007. Does mouse size affect study and evaluation results?: a study comparing preschool children's performance with small and regular-sized mice. In Proceedings of the 6th international Conference on interaction Design and Children (Aalborg, Denmark, June 06 - 08, 2007. IDC '07. ACM, New York, NY, 109-116.
- Hourcade, J.P., Bederson, B.B., Druin, A. & Guim-bretiere, F. 2004. Differences in Pointing Task Performance Between Preschool Children and Adults Using Mice. *ACM Transactions on Computer-Human Interaction*, 11(4), 357-386.
- Inkpen, K.M. 2001. Drag-and-Drop versus point-and-click: Mouse interaction styles for children. *ACM Transactions on Computer-Human Interaction* 8(1), 1-33
- Isbister, K., 2006. *Better Game Characters by Design – A psychological approach*. San Francisco, CA.: Morgan Kauffman.

- Isbister, K., Höök, K., Sharp, M., & Laaksolahti, J. 2006. *The sensual evaluation instrument: developing an affective evaluation tool*. Paper presented at the Proceedings of the SIGCHI conference on Human Factors in computing systems.
- Isen, A.M., 2000. *Positive Affect and Decision Making*. In : Handbook of emotion. Guilford, New York .
- Izard, C.E., 1972. *The face of emotion*. New York : Appleton-Century-Crofts.
- Izard, C., 1982, *Measuring Emotions in Infants and Children*, Cambridge: Cambridge University Press
- Izard, C. E., Libero, D. Z., Putnam, P., & Hayes, O. M. 1993. Stability of emotion experiences and their relation to traits of personality. *Journal of Personality and Social Psychology*, 64, 847-860.
- Johnson, C.N. 1988. Theory of mind and the structure of conscious experience. In J.W. Astington, P.L. Harris, and D.R. Olson (eds.), *Developing Theories of Mind*. New York: Cambridge University Press.
- Johnson, W. L., Soloway, E. 1984. Intention-Based Diagnosis of Programming Errors. *Proceedings of the National Conference on Artificial Intelligence*, 162-168.
- Kay, J. 1995. The um Toolkit for Cooperative User Modelling. *User Modelling and User Adapted Interaction*. 4, 149-196
- Kay, J. 1999. A Scrutable User Modelling Shell for User-Adapted Interaction. Ph.D. Thesis, Basser Department of Computer Science, University of Sydney, Sydney, Australia.
- Kay, J. 2001. Learner Control. User Model. *User Modelling and User Adapted Interaction* 11(1-2): 111-127
- Kay,J. & L. Li 2006): Scrutable Learner Modelling and Learner Reflection in Student Self-assessment. *Intelligent Tutoring Systems 2006*: 763-765
- Kelly, S.R., E. Mazzone , M. Horton , J.C. Read, 2006., Bluebells: a design method for child-centred product development, *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*, p.361-368, October 14-18, 2006, Oslo, Norway
- Kerly, A. & Bull, S. 2006. The Potential for Chatbots in Negotiated Learner Modelling, in M. Ikeda, K. Ashley & T-W. Chan (eds), *Intelligent Tutoring Systems: 8th International Conference*, Springer-Verlag, Berlin Heidelberg, 443-452.
- Kerr, R. 1975. Movement control and maturation in elementary-grade children. *Perceptual and Motor Skills* 41, 151-154.
- Klein, J., Moon, Y., and Picard, R., 2002. "This computer responds to frustration". *Interacting with computers*, 14(2), 2002, 119-140.
- Kok, A.J., 1991. A review and Synthesis of User Modelling in Intelligent Systems, *The Knowledge Engineering Review*, vol. 6, no. 1, pp. 21-47.
- Kort, B., Reilly, R., and Picard, R.W.,2001, "An Affective Model of Interplay Between Emotions and Learning: Reengineering Educational Pedagogy-Building a Learning Companion," In *Proceedings of International Conference on Advanced Learning Technologies (ICALT 2001)*, August 2001, Madison, WI.
- Krasner, G.E. & S.T., Pope 1988. A cookbook for using the model-view-controller user interface paradigm in Smalltalk-80. *JOOP*, 1(3), August, 1988.
- Kyriacou, C., 1997. *Effective Teaching In Schools: Theory and Practice*, 2nd Edition. Cheltenham: Stanley Thornes.
- Lang, P. J. *The Cognitive Psychophysiology of Emotion: Anxiety and the Anxiety Disorders*. Hillsdale, NJ: Lawrence Erlbaum, 1985.
- Lazarus, R.S., 1991, *Emotion and adaptation*, 1991, New York: Oxford University Press.
- LeDoux, J., 1996. *The Emotional Brain*. New York: USA: Simon and Schuster.
- Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. 1997. *The persona effect: affective impact of animated pedagogical agents*. Paper

- presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.
- Liebeck, P., 1984. *How Children Learn Mathematics: A Guide for Parents and Teachers*: England:Penguin Books.
- Linnenbringk, E.A. & Pintrich, P.R. 2002a. Achievement goal theory and affect: an asymmetrical bidirectional model. *Educational Psychologist*, 37, 69-78.
- Linnenbringk, E.A. & Pintrich, P.R. 2002b. The tole of motivational beliefs in conceptual change. In M.Limon & L. Mason (Eds.) *Reconsidering conceptual change: Issues in tehory and practice*. Dordrecht, The Netherlands: Kluwer Academic Publishers. pp. 115-135
- Linnenbringk, E.A. & Pintrich, P.R. 2003. Motivation, Affect, and cognitive processing: what role does affect play? Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Linnenbringk, E.A. & Pintrich, P.R. 2004. Role of affect in cognitive processing in accademic contexts. In D. Dai & R. Sternberg (Eds.). *Motivation, emotion, and cognition: Integrative perspectives on intellectual fonctionning and development* Mahwah, NJ: Lawrence Editions. pp. 57-87.
- Luckin, R., & Hammerton, L. 2002. Getting to Know Me: Helping Learners Understand Their Own Learning Needs through Metacognitive Scaffolding *Intelligent Tutoring Systems* (pp. 759-771. Berlin: Springer.
- Lloyd, T. & Bull, S. (2006). A Haptic Learner Model, *International Journal of Continuing Engineering Education and Lifelong Learning* 16(1-2), 137-149.
- Lorr, M., McNair, D.M. & Fisher, S. 1982: Evidence for bipolar mood states. *Journal of personality assessment*, 46, 432-436.
- Mabbott, A. & Bull, S. (2004). Alternative Views on Knowledge: Presentation of Open Learner Models, in J.C. Lester, R.M. Vicari & F. Paraguacu (eds), *Intelligent Tutoring Systems: 7th International Conference*, Springer-Verlag, Berlin Heidelberg, 689-698.
- Mabbott, A. & Bull, S. 2006. Student Preferences for Editing, Persuading and Negotiating the Open Learner Model, in M. Ikeda, K. Ashley & T-W. Chan (eds), *Intelligen Tutoring Systems: 8th International Conference*, Springer-Verlag, Berlin Heidelberg, 481-490.
- MacFarlane, S., & Pasiali, A. 2005. Adapting the Heuristic Evaluation Method for Use with Children. Paper presented at the Interact
- Malone, T. W., & Lepper, M. R. 1987. Making learning fun: a taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, Learning and Instruction. Volume 3: Conative and Affective Process Analysis* Hillsdale, NJ: Lawrence Erlbaum. pp. 223–253.
- Mandler, G., 1984. *"Mind and Body: Psychology of emotion and stress."* New York; Norton.
- Markopoulos, P., & Bekker, M. 2003. Interaction Design and Children. *Interacting with Computers*, 15, 141-149.
- Markopoulos, P., Read, J., MacFarlane, S., and Hoysniemi, J., 2008. *Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers (Interactive Technologies*. Morgan Kaufmann Publishers Inc.
- Mazza, R. & Dimitrova, V. (2004). Visualising Student Tracking Data to Support Instructors in Web-Based Distance Education, 13th International World Wide Web Conference - Alternate Educational Track, 154-161.
- Mazzone, E., & Read, J. 2005. Not just bits of paper – design sessions with children. *Journal of Department of Computing, University of Central Lancashire, Preston, UK*(4.
- Mazzone, E., Read, JC, Beale, R., 2008. Design with and for disaffected teenagers In Proc of Nordichi 2008. ACM Press, Lund, Sweden.

- McCalla, G., Vassileva, J., Greer, J. & Bull, S. 2000. Active Learner Modelling. In Gauthier, G., Frasson, C., and VanLehn, K. (eds.. Proceedings of the 5th International Conference on Intelligent Tutoring Systems, ITS 2000, 53-62.
- McCloud, S., 1993. *Understanding Comics: The Invisible Art*. New York, NY: Harper Perennial.
- McCloud, S., 2000. *Reinventing Comics: How imagination and technology are revolutionizing an Art Form*. New York, NY: Harper Perennial.
- McCloud, S., 2006. *Making Comics: storytelling secrets of comics, manga and graphic novel*. New York, NY: Harper Perennial.
- McQuiggan, S., Lester, J.. 2006. Diagnosing Self-Efficacy in Intelligent Tutoring Systems: An Empirical Study. In Ikeda, M., Ashley, K., Chan, T. (Eds.) ITS 2006. LNCS, vol 4053, pp. 565-574. Heidelberg: Springer.
- McQuiggan, S., Mott, B., & Lester, J.: Modeling self-efficacy in intelligent tutoring systems: an inductive approach. *User Modeling and User-Adapted Interaction*, 18(1-2), 81--123 2008.
- Meddis, R. 1972, Bipolar factors in mood adjective checklists. *British journal of social and clinical psychology*, 11, 178-184.
- Mehrabian, A. Basic Dimensions for a General Psychological Theory. OG&H Publishers. 1980.
- Mehrabian, A., and J. Russell "Evidence for a Three-Factor Theory of Emotions." *journal of Research in Personality* 11 (1977): 273-94.
- Meyer, M., and J.C. Turner, 2007. Scaffolding emotions in classrooms. In P.A. Schutz & P. Pekrun (Eds.), *Emotion in education* San Diego: Academic Press pp. 243-258.
- Mitrovic, A. & Martin, B. (2002). Evaluating the Effects of Open Student Models on learning, P. de Bra, P. Brusilovsky & R. Conejo (eds), *Proceedings of 2nd International Conference on Adaptive Hypermedia and Adaptive Web-based Systems*, Springer-Verlag, Berlin Heidelberg, 296-305.
- Mitrovic, A., and Martin, B., 2007. Evaluating the Effect of Open Student Models on Self-Assessment. *I. J. Artificial Intelligence in Education* 17(2): 121-144
- Morales, R., Pain, H., & Conlon, T. 2001. Effects of Inspecting Learner Models on Learners' Abilities. *Proceedings of AIED'01*. IOS Press, pp. 434-445.
- Moreno, R., Mayer, R., Spires, H., & Lester, J. 2001. The case for social agency in computer-based teaching: do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, 19, 177-213.
- Muehlenbrock, M., Tewissen, F. and Hoppe, H.U. 1998. A framework system for intelligent support in open distributed learning environments. *International Journal of Artificial Intelligence in Education*. 9, 256-274.
- Muller, M. J. 1991. *PICTIVE—an exploration in participatory design*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology.
- Nielsen, C. M., Overgaard, M., Pedersen, M. B., & Stage, J. 2005. Feedback from Usability Evaluation to User Interface Design: Are Usability Reports Any Good? *Human-Computer Interaction - INTERACT 2005* (pp. 391-404).
- Nielsen, C. M., Overgaard, M., Pedersen, M. B., & Stage, J. 2005. Feedback from Usability Evaluation to User Interface Design: Are Usability Reports Any Good? *Human-Computer Interaction - INTERACT 2005* (pp. 391-404).
- Nierenberg, G. I., 1977. *Fundamentals of Negotiating*. New York: Hawthorn Books.
- Norman, D. A., 1998. *Design of Everyday Things*: MIT Press.

- Nowak, K. & Rauh, C.: Choose your 'buddy icon' carefully: the influence of avatar androgyny, anthropomorphism and credibility in online interactions. *Computers Mediated Communication*, 11(1), article 8. 2008.
- Ortony, A., Clore, G.L., Collins, A., 1988: "*The cognitive structure of emotions*". Cambridge: Cambridge University Press
- Osgood, C.E. Dimensionality of the semantic space for communication via facial expressions. *Scandinavian Journal of Psychology*, 7(1), 1-30, 1966.
- Pantic, M. & Rothkrantz, L.J., 1997, "Automatic Recognition of Facial Expression and Human Emotions", ASCI'97 Proceedings of the third annual conference of the Advanced School for Computing and Imaging, Heijen, The Netherlands, June 2 – 4.
- Pekrun, R. Goetz, T., Titz, W., & Perry, R.P. 2002a. Academic emotions in students' self-regulated learning and achievement: A program of quantitative and qualitative research. *Educational Psychologist*, 37, 91-106.
- Pekrun, R., Goetz, T., Titz, W., & Perry, R.P. 2002b. Positive emotions in education. In E. Frydenberg (Ed.), *Beyond coping: Meeting goals, visions and challenges*. Oxford, UK: Elsevier. pp. 149-174
- Piaget, J., 1970. *Science of Education and the psychology of the child*. New York: Orion Press.
- Piaget, J., 1974. *To understand is to invent: The future of education*. New-york : Viking Press.
- Picard, R.W., 1997., "Affective Computing," MIT Press, Cambridge, 1997.
- Pintrich, P.R. 2000. The role of goal orientation in self-regulated learning. In M. Boekarts, P.R., Pintrich & M. Zeider (Eds.), *Handbook of self-regulation: Theory, research and applications*, San Diego: Academic Press. pp. 451-502
- Pons, F., Lawson, J., Harris, P. L., & De Rosnay, M. 2003. Individual differences in children's emotion understanding: effects of age and language. *Scandinavian Journal of Psychology*, 44, 347-353.
- Preece, J., Rogers, Y., & Sharp, H., 2002. *Interaction design beyond human-computer interaction*, New York: Wiley.
- Prendinger, H., Ishizuja, M., 2005. The Empathic Companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence* 19(3-4), 267-285.
- Read, J.C., M. Horton, and E. Mazzone, 2005. The Design of Digital Tools for the Primary Writing Classroom In Proc of EdMedia'05Montreal, CA, 2005.
- Read, JC, MacFarlane, SJ, and Casey, C. 2002. En-durability, engagement, and expectations: Measuring children's fun. Proceedings of the Conference on Interaction Design for Children (pp.189-198). Eindhoven: Shaker Publishing.
- Read, J.C. 2008. Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. *Cognition Technology and Work*. 10, 2 (2008), 119-128.
- Read, J.C. and MacFarlane, S. 2006. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. Proceedings of the 2006 conference on Interaction design and children (Tampere, Finland, 2006), 81-88.
- Renninger, K.A., 2000: Individual interest and its implications for understanding intrinsic motivation. In C. Sandone & J.M. Harackiewicz (Eds.) *Intrinsic and extrinsic motivation: the search for optimal motivation and performance* (pp. 373-404). San Diego, CA: Academic Press.
- Richmond, P. G., 1970. *An Introduction to Piaget*, London: Routledge & Kegan Paul.
- Robertson, J., 2000. Is Attribution Training a Worthwhile Classroom Intervention for K-12 Students with learning difficulties? *Educational Psychology Review* 12(1), 111-134

- Rode, J. A., Stringer, M., Toye, E. F., Simpson, A. R., & Blackwell, A. F. 2003. *Curriculum-focused design*. Paper presented at the Proceedings of the 2003 conference on Interaction design and children.
- Ruch, W., & Ekman, P. (2001). The expressive pattern of laughter. In A. Kaszniak (Ed.), *Emotion, qualia and consciousness* (pp. 426-443). Tokyo.
- Saarni, C. 1988. Children's understanding of the interpersonal consequences of dissemblance of nonverbal emotional expressive behavior. *Journal of Nonverbal Behavior*, 12, 275-294.
- Salovey, P., Bedell, B., Detweiler, J., & Mayer, J. D. 2000. Current directions in emotional intelligence research. In M. Levis & J. M. Havilland-Jones (Eds.), *Handbook of emotions*. New York: Guilford Press. pp. 504-520
- Sander, D., & Scherer, K.R., 2009. *The oxford companion to emotion and the affective sciences*. Oxford: oxford university press.
- Scherer, K.R. and Peper, M. (2001). Psychological theories of emotion and neuropsychological research. In: F. Boller and J. Grafman (eds), *Handbook of neuropsychology*, Vol. 5, pp. 17-48). Amsterdam: Elsevier.
- Schneider, K.G., 1996. Children and information visualization technologies. *Interactions*. 3(5), 68-73.
- Schlosberg, H. S. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology*, 44, 229-237.
- Schwarz, N. & Clore, G. L., Mood, Misattribution, and Judgements of Well-Being: Informative and Directive Functions of Affective States. *JPSP*, 1983,45,513-523
- Self, J. 1988. Bypassing the intractable problem of student modelling. In Gauthier, G., and Frasson, C., (eds.), *Proceedings of Intelligent Tutoring Systems '88*, 18-24. Self, J. 1994. Formal Approaches for Student Modeling. In Greer, J., McCalla, G. (eds.), *Student Modelling: the Key to Individualized Knowledge-based Instruction*. 295- 352.
- Self, J., 1999. The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. *International Journal of Artificial Intelligence in Education*. 10, 350- 364.
- Shahrour, G. & Bull, S. (2009). Interaction Preferences and Learning in an Inspectable Learner Model for Language, in V. Dimitrova, R. Mizoguchi, B. du Boulay & A. Graesser (eds), *Artificial Intelligence in Education 2009*, IOS Press, Amsterdam, 659-661.
- Shneiderman, B., 1998. *Designing the user interface: strategies for effective human-computer interaction*: Third Edition. Reading, Massachusetts: Addison-Wesley.
- Shneiderman, B., 2004. Designing for fun: how can we design user interfaces to be more fun? *Interactions*, 11(5), 48-50.
- Shute, V. J., & Psotka, J. 1996. Intelligent tutoring systems: Past, present, and future. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* New York, NY: Macmillan. pp. 570-600.
- Shute, V. J. & Zapata-Rivera, D. 2010. Educational measurement and intelligent systems. In E. Baker, B. McGaw, & P. Peterson (Eds.), *Third Edition of the International Encyclopedia of Education*. Oxford, UK: Elsevier Publishers.
- Sleeman, D., Smith, M. J. 1981. Modelling Student's Problem Solving. *Artificial Intelligence*, 16, 171-187.
- Stein, N., Trabasso, T., & Liwag, M. 2000. A goal appraisal theory of emotional understanding: Implications for development and learning. In M. Lewis & J. Haviland (Eds.), *Handbook of emotions, 2nd Edition*. New York: Guilford Press. pp. 436-457
- Straker, A, 1989. *Children using computer*, Blackwell
- Sylva, K., & Lunt, I., 1982. *Child Development A First Course*. Oxford: Blackwell.

- Takeuchi, A., & Naito, T. 1995. *Situated facial displays: towards social interaction*. Paper presented at the Human Factors in Computing Systems: CHI'95 Conference Proceedings.
- Thomas, F., & Johnson, O. (1981). Disney animation: the illusion of life: Walt Disney Productions.
- Thrash, T.M., & Elliot, A.J. (2004). Inspiration: Core characteristics, component processes, antecedents, and function. *Journal of Personality and Social Psychology*, 87, 957-973.
- Tricot, A., Plégat-Soutjis, F., Camps, J.-F., Amiel, A., Lutz, G., Morcillo, A. « Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH » *Environnements Informatiques pour l'Apprentissage Humain 2003*, France (2003)
- Tronick, E., & Morelli, G. 1991. Foreword. In B. L. J. Nugent & Brazelton (Eds.), *The cultural context of infancy: Vol2. Multicultural and interdisciplinary approach to parent-infant relations*. Norwood: NJ:Aplex. pp. ix-xiii
- Tudor, L. G., Muller, M. J., & Dayton, T. 1993. *A C.A.R.D. game for participatory task analysis and redesign: macroscopic complement to PICTIVE*. Paper presented at the INTERACT '93 and CHI '93 Conference Companion on Human Factors in Computing Systems.
- L.Uden and A. Dix (2000). Iconic Interfaces For Kids On The Internet. IFIP World Computer Congress, Beijing, pp. 279-286
- Van Dam, A., Becker, S., & Simpson, R. M., 2005. Next-generation educational software. *EDUCAUSE review*, March/April 2005, 27-42.
- Van den Brande, L., 1993. *Flexible and Distance Learning*. Chichester: Wiley
- Vincent, C., Delozanne, E., Grugeon, B., Gelis, J.-M., Rogalski, J., & Coulange, L. 2005. *Des erreurs aux stéréotypes: des modèles cognitifs de différents niveaux dans le projet Pépité*. Paper presented at the Actes de la conférence EIAH 2005, Montpellier, 25-27 mai 2005.
- Vorderer, P. and Bryant, J. 2006. *Playing video games: motives, responses, and consequences*. Routledge.
- Vygotsky, L. S., 1987. *The collected works of L. S. Vygotsky: Volume I: Problems of general psychology*, New York: Plenum Press.
- Walker, J., Sproull, L., & Subramani, R. 1994. *Using a human face in an interface*. Paper presented at the CHI 1994 Conference on Human Factors in Computing Systems.
- Weber, G. & Brusilovsky, P., 2001. EML-ART: An Adaptive Versatile System for Web-Based Instruction, *International Journal of Intelligence in Education*, 12(4), 351-384.
- Wehmeyer, M. L., Agran, M., & Hughes, C., 2000. A national survey of teachers: promotion of selfdetermination and student-directed learning. *The Journal of Special Education*, 34, 58-68.
- Weiner, B. 1972. Attribution Theory, Achievement Motivation, and the educational. *Review of Educational research*, 42(2), 203-215.
- Weiner, B., 2007. Examining emotional diversity in the classroom: An attribution theorist considers the moral emotions. In P.A. Schutz & P. Pekrun (Eds.), *Emotion in education*. San Diego:Academic Press pp. 55-84
- Williams, M.R., Cross, D.I., Hong, J.Y., Aultman, L.P., Osbon, J.N., & Shutz, P.A. 2006. "There are no emotions in math": How teachers approach emotions in the classroom. *eachers College Record Volume 110 Number 8, 2008*, p. 1574-1610
<http://www.tcrecord.org> ID Number: 15153, Date Accessed: 10/6/2010

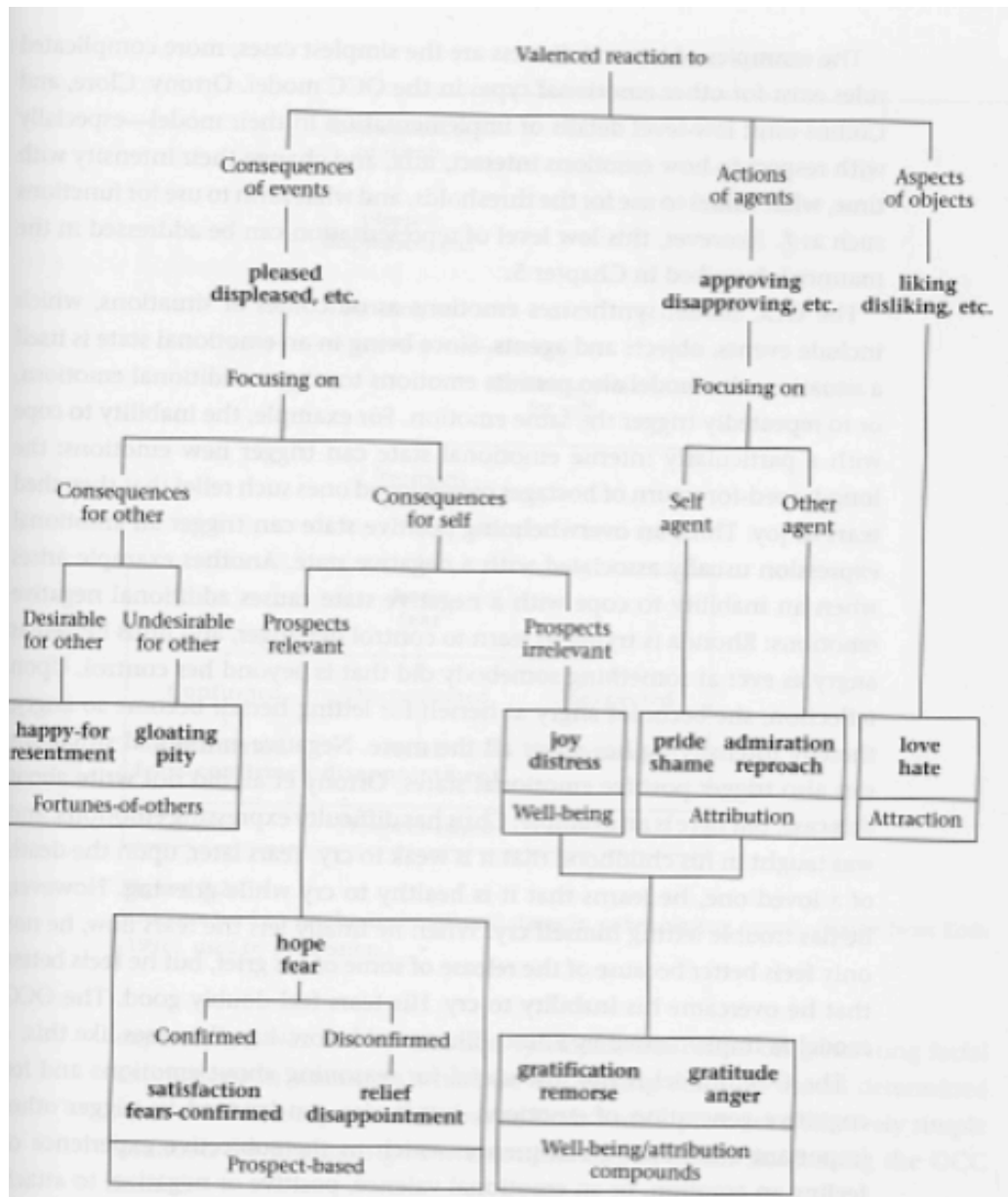
- Wood, D., Bruner, J.S., & Ross, G., 1976. The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry*, 17, 89-100.
- Wood, H., & Wood, D., 1999. Help Seeking, Learning and Contingent Tutoring. *Computer and Education*, 33, 153-169.
- Woolf, B.P., Burleson, W., Arroyo, I., Dragon, T., Picard, R. 2009): Emotional Intelligence for Computer Tutors. In El Kaliouby, R., Craig, S. (eds.) Special Issue on Modelling and Scaffolding Affective Experiences to Impact Learning, *International Journal of Learning Technology*.
- Woolf, B.P., Arroyo, I., Muldner, K., Burleson, W., Cooper, D., Dolan, R., Christopherson, R., 2010a), The Effect of Motivational Learning Companions on Low Achieving Students and Students with Disabilities, In V. Aleven, J. Kay and J. Mostow (Eds.) *Intelligent Tutoring Systems: 10th International Conference, ITS 2010* (pp. 327-337). LNCS 6094. Berlin Heidelberg: Springer-Verlag.
- Woolf, B.P., I.Arroyo, D. Cooper, W. Burleson, and K. Muldner, 2010b. Affective Tutors: Automatic Detection of and Response to Student Emotion, *In: R.Nkambou et al* (Eds.): *Advances in Intelligent Tutoring Systems*. SCI 208, Berlin: Springer Verlag Heidelberg, pp207-227
- Wyeth, P., & Wyeth, G. 2001. *Electronic Blocks: Tangible Programming Elements for Preschoolers*. Paper presented at the Human-Computer Interaction - INTERACT'01.
- Yacef, K. 2005. The Logic-ITA in the classroom: a medium scale experiment. *International Journal of Artificial Intelligence in Education*, 15(1), 41-60.
- Yee, N. & Bailenson, J., 2007. The Proteus Effect: Self-transformations in virtual reality. *Human Communication Research*, 33, 271—290.
- Zapata-Rivera, J.D. 2002. cbCPT: Knowledge Engineering for CPTs in Bayesian Networks. *In Proceedings of the Fifteenth Canadian Conference on Artificial Intelligence AI 2002*, 368- 370.
- Zapata-Rivera, J.D 2003) Learning Environments based on Inspectable Student Models. Ph.D. Thesis. University of Saskatchewan. Department of Computer Science.
- Zapata-Rivera, J.D., and Greer, J., 2004. Inspectable Bayesian Student Modelling Servers in Multi-Agent Tutoring Systems. *International Journal of Human Computer Studies*. 61(4). 535-563.
- Zeidner, M. 1998. *Test Anxiety: the state of the art*. New York: Plenum.
- Zeng, Z. et al, 2009. A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *Patterns Analysis and Machine Intelligence*. IEEE Transactions on 31(1): 39-58, 2009.
- Zimmerman, B., 2000. Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*, 25, 82-91.

APPENDICES

APPENDIX A: LITERATURE REVIEW	255
APPENDIX B: CHILDREN'S UNDERSTANDING OF THE METAPHORS	257
APPENDIX C: EMBODIED PEDAGOGICAL AGENTS	260
APPENDIX D: ANALYZING CHILDREN'S INTERACTION STYLES IN EDUCATIONAL SOFTWARE	282
APPENDIX E: HOW WELL DO PEOPLE UNDERSTAND AND CLASSIFY EMOTIONS?	291
APPENDIX F: EMOTION SETS FOR PEDAGOGICAL SOFTWARE WITH CHILDREN AND TEACHERS	299
APPENDIX G: ASSESSING INTERFACE CHARACTER EMOTIONAL STATE	306
APPENDIX H: ASSESSING YOUR OWN EMOTIONAL STATE, USING SOREMO FOR THE FIRST TIME	312
APPENDIX I: OLM REPRESENTATION	322
APPENDIX J: SMILEY FACES, TRAFFIC-LIGHTS, OR BOTH IN THE USER MODEL	329

APPENDIX A: LITERATURE REVIEW

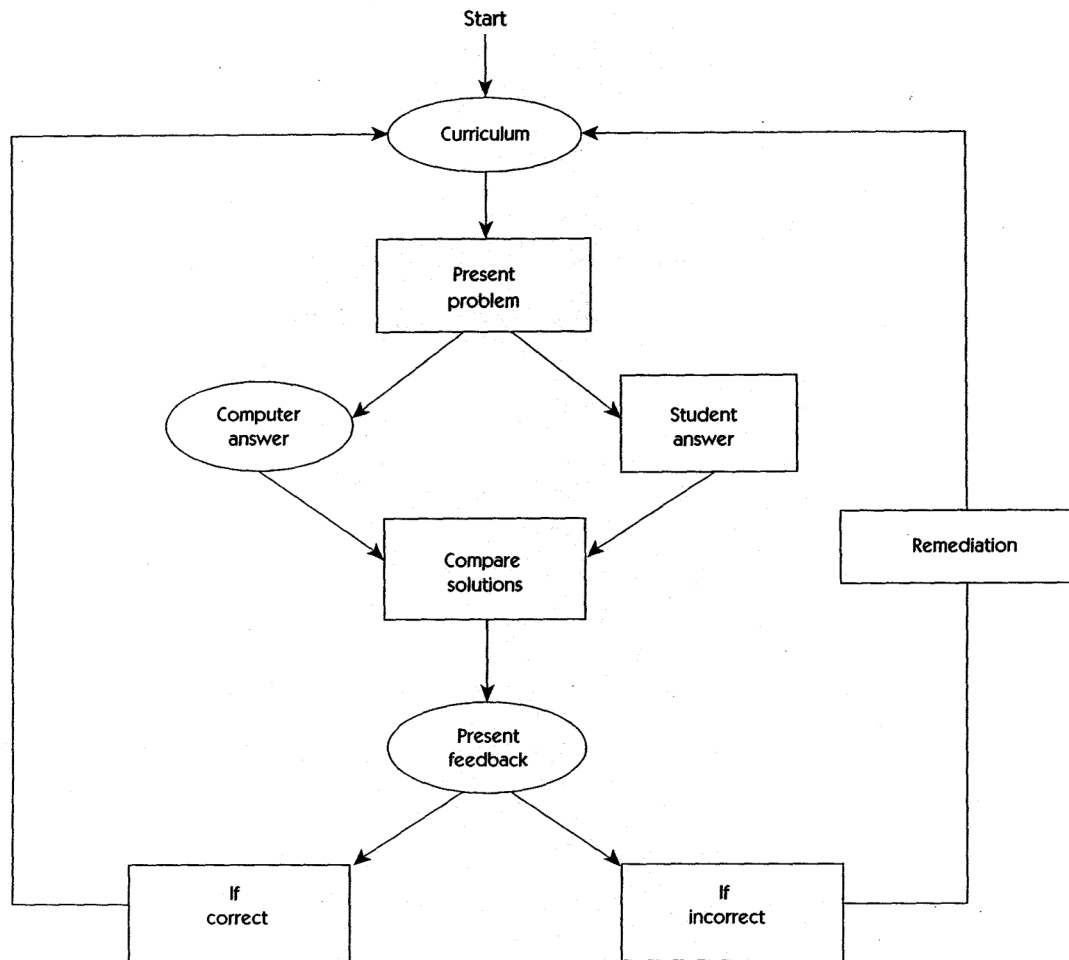
THE OCC MODEL



The OCC model structure: Ortony, Clore & Collins, 1988.

DETAILED STRUCTURE OF THE INTERACTION WITHIN ITS.

COMPUTER-ASSISTED INSTRUCTIONS



Extracted from (Shutte & Psotka, 1996)., p572.

APPENDIX B: CHILDREN'S UNDERSTANDING OF THE METAPHORS

STUDY AIM

Empirical studies were conducted on primary school children from French and English schools investigating their understanding of these two metaphors of learning when used in pedagogical context: the Traffic-Light system and the Smiley Faces metaphors.

PARTICIPANTS

Two primary school classes participated to the study: a French class of twenty-nine children (average eight yrs), and an English class of twenty-six children (average ten yrs). The educational products designed in this PhD aim at being tested on both French and English primary school children. For this reason, the use of the symbols representing the metaphors need to be usable by international users, and not culturally dictated. Therefore both cultures were represented in the users participating in the study.

PROCEDURE

Each child was presented with an array containing three learning situations. Each situation represents the achievements of a child in the learning of multiplication tables in the form of an affirmative sentence.

S/he had also in her/his possession a picture of each state of learning for both metaphors (three Traffic-Lights, and three Smileys), and their meaning in words ('I fully understand', 'I need more practice', 'I do not understand yet').

They were asked to read carefully each sentence, and to select one picture of Smileys, another of traffic-lights, and at last a sentence that represented what the teacher would give the student as a feedback in the situation. The children needed to use the pictures and select the one that they considered best represented the level of the child taken as an example. They therefore needed to associate the various representations with the understanding of the pedagogical context.

In order to avoid children not being able to realize the task due to learning difficulties, each sentence was read to the class by the teacher, giving children the time to answer before concentrating on the next sentence. The whole sentence was repeated three times, the children

were allowed to change their selection. Only the understanding of the three levels of learning feedback (fully/partial/not understanding) and their association to the metaphors of learning were investigated here.

At the end, the children were asked to confirm their answer by gluing the pictorial representations and the written sentences in the appropriate calls of the array. The material was presented in French to the French class of participants, and in English to the English class of participants.

An imaginary example of participant's ratings is illustrated in Figure B.1:






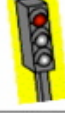
Smiley	Traffic-Light	Level	Multiplication Results
		I fully understand	"Gwenaël only answered correctly 4 multiplications out of 10 on the table of 3."
		I need more practice	"Anthony got 10 out of 10 on the table of 9."
		I do not understand yet	"Sasha tried the table of 2, it was all wrong. »"

Figure B.1: Test grid for the understanding of metaphors

The child would have correctly identified the different levels of learning when using the graphical representations (Smileys and Traffic-Lights). However, for the written metaphors, the two first sentences would be out of order.

RESULTS

The results of the study are presented in table B.1:

Table B.1: Understanding of the metaphor by system (Percentage)

Class	Nb	Age	Traffic-Light	Smiley Faces	Sentences
French Class	29	8	93%	96.5%	100%
English Class	28	10	96%	100%	100%
TOTAL	57	9.5	94.5%	98.2%	100%

The concept of the levels of learning is fully grasped as all children in both class associated the right sentences to the learning situations.

Whilst the understanding of the Traffic-Light (TL) metaphor is slightly lower than the Smiley Faces for both classes, they are both well associated with the learning states with an association rate higher than or equal to 93% in all cases. The misidentifications of Traffic-Lights mainly concerned the middle answer, with children willing to give the middle state (orange TL) instead of the lowest (red TL) to Sasha, and higher state (green TL) instead of middle one (orange TL) to Gwenaël. In this case, they did not use the red TL and drew another green TL for the situation portrayed by Anthony (high level of achievement). The children were asked to complete the task by giving answers as the teachers would, in order to reduce the overestimation of levels out of child's sympathy for their imaginary peers.

Using the Traffic-Light metaphor in isolation can lead to visual problems for colour-blind children who may not be able to distinguish between green and red. Whilst the appropriate context and position of the colours on a Traffic Light helps users distinguish which colour is represented, the use of the colours outside the Traffic-Lights context can cause difficulties of understanding. This can be overcome by combining the two metaphors in the development of interface elements. The redundancy of symbolism (i.e. the use of both colour and expression) should therefore enable children to bypass any handicap linked to the perception of colours.

From these results above, we consider both metaphors well enough understood to include them in the design of educational components. A second series of task coloured Smileys (red – sad, orange – neutral, green – happy) was conducted with French children from two classes (10 aged 7-9 yrs, 10 aged 9-11 yrs) following a similar procedure, and considering a combination of both metaphors (mixing colours and facial expressions) rather than each metaphor as a separate entity. The results showed that all children (100%) correctly associated the pictorial representations with the learning situations.

APPENDIX C: EMBODIED PEDAGOGICAL AGENTS

Usually embedded within the software to aid social and communicative features, (Johnson, 2003; McQuiggan et al, 2008), EPAs can be found in educational software in increasing numbers. They differ in visual representations, communication styles, or pedagogical roles. EPAs can adopt one or more different pedagogical roles within the same application. Children's response to them vary in their willingness to follow the pedagogical behaviours embedded into the agent, their capacity to make a transfer of learning occur (Baylor & Kim, 2005), and children's own self-confidence in answering questions on some subjects (Baylor, 2005).

Research on the user interface design of agents, and in the design of comics over the past twenty years, has focused on the impact of the degree of detail and naturalness of the EPA. These factors affect their processing and interpretation by users (Cook, 1979; Isbister, 2006), in addition to self-identification processes, and engagement level with the EPA (Gulz & Haake 2006; McCloud, 1993). Other studies have investigated the impact of 'instructional roles' on learning and motivation for particular age groups (Baylor & Kim, 2005). A range of pedagogical roles used in EPAs worldwide, have been classified in Haake & Gulz (2009) on the dimension of authority. In Haake & Gulz (2009), three EPA factors were studied with children aged 12-15: visual static appearance, communicative style, and pedagogical role. Some interesting and potentially unexpected results were found, such as when female students chose 'learning companions' they preferred more stylized, visual characters. However, as children grow, their interests change. This impacts their ability (Piaget, 1970) and willingness (Acuff and Reiher, 1997) to use intelligent environments.

The literature however lacks results on the role of EPA's look and pedagogical roles concerning our users' age range. Section C.1 introduces the results of an exploratory study on children's preference in EPAs performed during the design of our OLM software. Section C.2 presents a design study using children as design informants in order to create a new EPA for the *Multiplotest* application. Section C.3 details the design decisions made from the results of the study and the corresponding literature on the design of EPAs for educational software.

C.1 WHAT DO CHILDREN FAVOUR AS EPA?

In this section, a study is introduced, where 86 primary school children (aged 7-11) chose an EPA to ‘accompany’ them in their learning of multiplications in the ITS application, *Multiplotest*.

PRELIMINARY PARTICIPATORY DESIGN

When investigating the impact of visual characteristics of digital characters, and the agent’s behaviour, on children’s choices of EPA, a first step is to define the visual dimensions to be taken into account, and the role the EPA will have to hold in the interface.

Dimensions of EPA’s design

In Haake & Gulz (2009), a theoretical framework is defined to evaluate user’s EPA choice along three design factors: visual static appearance, communicative style, and pedagogical role. In this work on EPAs, a choice was made to evaluate the impact of two components of visual static appearance (degree of detail, degree of naturalism) in relation to the EPA’s pedagogical role and the communicative style. The communicative style was decided in participatory-session with the teachers, according to the pedagogical goals defined. The results enable comparison with Haake & Gulz’s study as to the user’s preferences, according to the age of the child-participants.

One contribution of this study is that the user-group targetted differ from the Haake & Gulz’s study in culture, and more especially in age: they do not belong to the same Piagetan stage of cognitive development (Piaget, 1970), nor to the same stage in Acuff and Reiher’s categorization (1997). Children aged 11-15 years relate more strongly to more realistic characters, preferring realistic to fantasy worlds (Acuff & Reiher, 1997), and are able to use abstract thinking to solve problems (Piaget, 1970). According to Piaget (1970), children aged 7-11 (within the concrete operational stage) can think logically, but not abstractly. Whilst they can distinguish between reality and fantasy, they are developing beyond a stage where perception is dominant, and thus potentially misled by what they see. According to Acuff & Reiher (1997), the youngest of our age group participants are leaving the ‘emerging/autonomy’ stage to enter the ‘rule/role’ one. They are no longer dominated by a

world of fantasy and magic, where there was a need of stimulation associated with comfort and love. By the age of 8, their interest shifts gradually from fantasy to reality.

For this reason, when considering the degree of naturalism, a choice was made to investigate the impact of the level of anthropomorphism in the design versus a character based on a smiley face. This should help gauge children's interest in interacting with an EPA that is more fantasy-like (smiley-shaped), or more redolent of the real world (humanoid-shaped). The use of children and teachers as participatory-design partners in the design of the EPA's visual appearances for the study will aid in producing a sample of EPAs comprised of graphical components both familiar and appealing to children of this age.

Defining the agent's visual appearance.

The format of the pictorial representations to be used in this study arose out of a participatory design session with 2 teachers and 20 children aged 7-11, who did not partake to the main study afterwards. Different children were used so as not to induce a bias in the experiment as the children's choice in their preference of EPAs could be influenced by the design activity itself. For example, a child could be more willing to select a picture designed by someone s/he likes, or refuse to choose another he would normally have chosen, just because he was not the one to select it during the process.

During the participatory design sessions the children first perused a collection of pictures taken from Internet picture databases on learning companions and the comic literature. All pictures had some emotional facial expression, and were all redesigned by a comic designer so as to portray the same emotion: happiness. The sample of pictures represented various levels of details, use of colours, 2D/3D visualisations, and represented human faces, animated objects, or animals. A first sample of characters was chosen by researchers from the lab when looking for online databases. Then, each child perused the sites found, adding their own links, and created a list of 10 characters they liked the most. Each child then designed their own instructor and a learning companion using pens, paints, potter, or 'Draw' software on the computer in the classroom.

All of the children chose and drew female characters as instructors. When interviewed, they explained that the picture should be as similar to their teacher as possible. The participating school consisted entirely of female teachers, and therefore only one type of humanoid character for the instructor will be proposed, with a female gender. However, when defining the learning companion, the children drew a mixture of female and male child characters, similar to their age. Therefore two characters will represent each naturalistic

rationale given during the participatory-design sessions was that the children felt they could associate with the characters quite well, and their underlying personalities represented quite accurately the need for guidance in mathematical domain that a learning companion could represent.

GOAL OF THE STUDY

The experimental study was performed with three French classes from a French primary school. The goal of the study was to investigate users' choice of EPA with respect to their visual appearance and pedagogical role. In particular, possible relationships between these variables were investigated with regards to user preferences.

PARTICIPANTS

86 children aged 7-11 (46 girls, 40 boys) from a French primary school participated in the study. The students came from three classes of two levels (one class of CE1, one class CE1/CE2, and one class of CE2). The majority of students had no familiarity with pedagogical agents comprised of embodied computer characters, but they were all familiar with the pictorial representation of naturalistic characters.

THE EPA'S SELECTION

Visual Appearance: Detailed vs. Simplified, Naturalistic vs. Stylized characters.

Figure C.2 illustrates the sets of characters the children had to choose from: the four characters grouped on the left representing the teacher/instructor characters, and the group on the right representing the learning companions.

Two approaches were taken to the design of the characters used in the study: varying the axis *degree of naturalism*, and the *degree of detail*.

Characters 1, 3, 6, and 8 made use of a stylized (smiley-shaped) representational form, while the other characters (2, 4, 5, 7, 9, 10) made use of a naturalistic one (humanoid-shaped).

The characters are again separated in terms of level of detail, with the top row 3D-rendered and detailed, and the bottom row 2D-rendered and simplified.

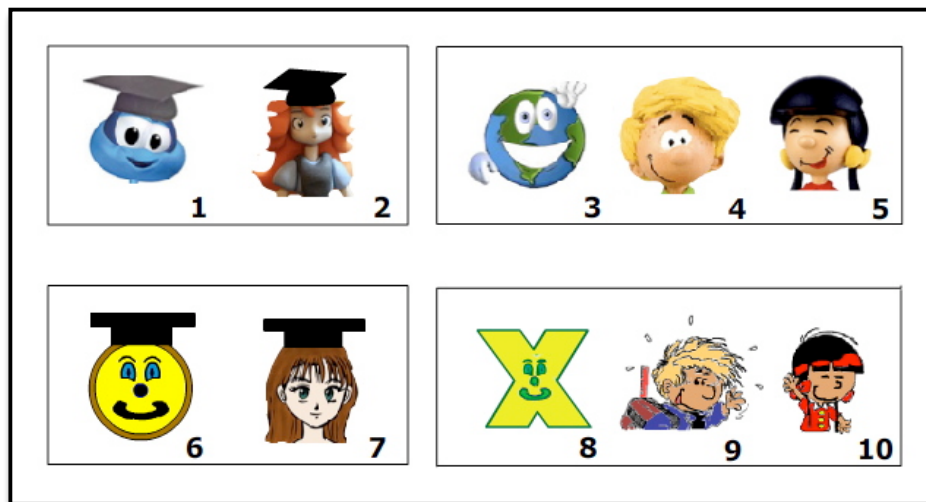


Figure C.2 The four sets of characters used in the study: each different in pedagogical goal and visual appearance.

Pedagogical Role: Instructor/Teacher vs. Learning Companion

From the categorization of EPA's roles in the literature (Haake & Gulz, 2009; Chou et al, 2003), and the studies undertaken on their impact on ITS (Baylor & Kim, 2005), the two following EPAs, different in pedagogical roles were investigated:

- An EPA representing an authoritative instructor, with an instructional role of 'expert' as defined in (Baylor & Kim, 2005), and mainly task-oriented as designed in (Bickmore, 2003).
- An EPA representing a non-authoritative, collaborative learning companion, with a 'mentor' instructional role, and a combined task-and-relation-oriented communicative style.

The two scenarios of use differed in the interaction with the EPA, but were equal in terms of pedagogical goal and activity architecture: the children could either choose one multiplication (level 1) or all multiplication tables (level 2) simultaneously to test their knowledge against. In this study, the impact of the EPA was investigated only on the use of mathematical activities, activities, that could be comparable in terms of user actions, if not level of difficulty. For this reason, children could choose to perform a test on the knowledge of multiplication tables without the misconceptions and method apprenticeship one has to learn in order to perform level 3, and level 4. Level 3 was authorized as the best students

might find level 1 too easy, if they already know all multiplication tables. The only restriction in the game was to choose a unique activity to perform within level 1 and level 2 activities. Each visual representation of the characters was investigated in the scenarios, the interface modulated to integrate their design.

EXPERIMENTAL DESIGN AND PROCEDURE

The participants were presented with both versions of the *Multipliote* ITS aimed at helping children learn multiplications: one version with an instructor, and another with a learning companion. The order in which they accessed software was counter-balanced (half of them beginning with the instructor and then the companion, and vice versa). The procedure for the study is illustrated in Figure C.3:

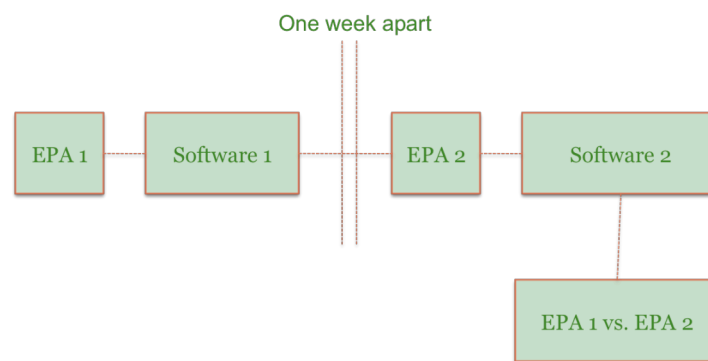


Figure C.3: Study Procedure

When in front of the software, the children were to choose a character as their EPA given the choices presented in Figure C.2, and then perform an activity within 5 minutes. They then followed the same procedure for the other condition. Finally, they were requested to choose the EPA they preferred, and explain why it was more appealing to them. In this experimental manipulation, the factors ‘pedagogical role’, ‘degree of detail’ and ‘degree of naturalism’ of the characters act as independent variables, and the user’s choice acts as the dependent variable.

RESEARCH HYPOTHESES

The following hypotheses were studied in this experiment:

Hypothesis 1

H₀¹: There is no difference in the number of learning companions versus instructors, preferred as an EPA.

H₁: Children will choose the learning companion more than the instructor as the software EPA.

Hypothesis 1 relates to the children's final choice, at the end of the session, after using both instructor and learning companion conditions. The hypothesis is concerned with the children's overall preference for either an instructor or learner. It is separated from the other hypotheses, and based on Acuff & Reiher's definition of a developmental stage (1997), where children still need support and comfort, and respond more positively to working with a digital peer than an instructor telling them what to do.

Hypothesis 2 and 3

H₀²: When introduced as a learning companion, there is no difference in the number of EPAs being chosen with a humanoid-shaped or smiley-shaped appearance.

H₂: The learning companion version will yield a preference for a smiley-shaped EPA.

H₀³: When introduced as an instructor, there is no difference in the number of EPAs being chosen with a shaped-shaped or smiley-shaped appearance.

H₃: The instructor version will yield a preference for a humanoid-shaped EPA.

Hypotheses 2 to 5 were drawn from theoretical work, field observations when working with our participatory-design partners in the definition of the pictorial representation, and children's reactions in usability studies with participants of the same age group in the design process of the *Multiplotest* software design.

Hypothesis 4 and 5

H₀⁴: When introduced as a learning companion, there is no difference in the number of EPAs being chosen with high or low levels of detail.

H₄: The learning companion version will yield a preference for an EPA with low levels of detail.

H₀⁵: When introduced as an instructor, there is no difference in the number of EPAs being chosen with high or low levels of detail.

H5: The instructor version will yield a preference for an EPA with low levels of detail.

H4 and H5 follow McCloud's theoretical framework (1993) where simplified characters amplify the meaning of an image, such as the character's affective components, and therefore afford more powerful social-emotional communications. This may help in the rational aspects of the EPA presented.

RESULTS

Hypothesis 1: Choice of Pedagogical Role for EPAs in ITS

Table C.1: Test data on the final choice between instructor and learning companion

Instructor	Companion	Total
21	65	86

A Chi-square 'goodness to fit' test demonstrates that there is a significant difference ($\chi^2=22.512$, $df=1$, $p < 0.001$) between the expected and observed frequencies, which rejects H_0^1 : children prefer EPAs in the role of learning companions when working on *Multiplicotest*.

The results from the study seem to show the same tendency of children favouring learning companions: in the number of pictures selected by children from the database a greater number was chosen as a representation of a learning companion than a teacher-like feature. Furthermore, when asked to select their favourite character at the end of the participatory design session, between the 10 proposed, 18/20 of the children chose the character they identified as best for learning companion. From these results, it can therefore be argued that children aged 7-11 might be more likely to choose a learning companion as EPA, independently of software considered.

Hypothesis 2 to 5: Associations of Ped. Role and Visual Static Appearance

The test data concerning hypotheses 2 to 5 can be presented in frequencies in a three-dimensional contingency table categorized by the variables: pedagogical role (P), visual style: degree of detail (D), and visual style: naturalism (N).

Table C.2: Test data categorized by the three variables: P, D, and N.

Pedagogical Role	Detail	Naturalism		Total
		<i>Naturalistic</i>	<i>Stylized</i>	
<i>Instructor</i>	<i>Detailed</i>	5	18	23
	<i>Simplified</i>	61	2	63
	Total	66	20	86
<i>Learning Companion</i>	<i>Detailed</i>	0	1	1
	<i>Simplified</i>	5	80	85
	Total	5	81	86
	Column Total	72	100	172

Hypotheses 2 and 3: Association of Pedagogical Role and Level of Naturalism

With the data separated into the two different pedagogical roles of instructor and companion, a Chi-square ‘goodness to fit’ test revealed a significant difference in the choice of humanoid-shaped/smiley-shaped visual appearance for both pedagogical roles.

Table C.3: χ^2 tests on the data of Table IV.3 concerning the choices of the level of naturalism in the pedagogical agents.

Condition	χ^2	df	p-value
Instructor	24.605	1	< 0.001
Learning Companion	67.163	1	< 0.001

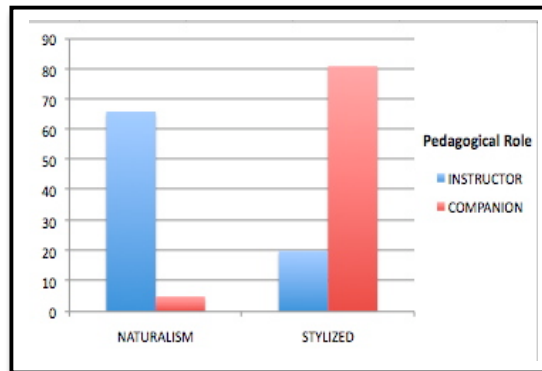


Figure C.3 Association of Pedagogical Role and Level of Naturalism

The null hypothesis H_0^2 and H_0^3 are rejected: The instructor version yields a preference for a naturalistic agent ($\chi^2=24.605$, $p < 0.001$), and the learning companions for a more stylized appearance ($\chi^2=67.163$, $p < 0.001$).

When asked the rationale for their choice in EPA, 60 out of the 86 children participating in the study claimed that the non-human simplified character (picture 8 in Figure IV.1) was better for this educational software. Children indeed identified with the character, and recognized it as the usual multiplication symbol, used in paper tests at school. In this character, surnamed “multiplication master” by some of the children, they saw a potential of gaining more guidance than the human characters: “Cedric is not good in Maths... he won’t be much help” said one of the children. The children preferring the more naturalistic representations of the EPA playing a teacher role expressed a willingness to get closer to the real world setting: “this character is a teacher, he should look like one”. Similar comments were given when considering the degree of details: the character number 2 in Figure IV.1 appeared too surreal for the application. With the character number 7, the children identified the EPA with their own classroom teacher, and found that the more detailed character looked too much like a “game-like” character.

Hypotheses 4 and 5: Association of Pedagogical Role and Degree of Details

With the data separated into the two different pedagogical roles of instructor and companion, a Chi-square ‘goodness to fit’ test revealed a significant difference in the degree of detail chosen in the visual appearance for both pedagogical roles. Figure C.4 and table C.3 illustrates that for the case of a learning companion, all but one child chose the simplified version of the companion, represented in Figure C.2 by the characters 8, 9, and 10.

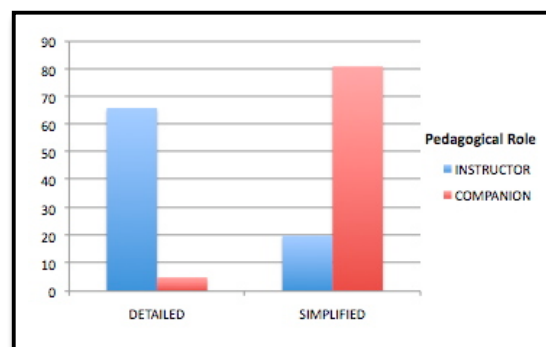


Figure C.4 Association of Pedagogical Role and Level of Details

Table C.5: χ^2 tests on the data of Table IV.2 concerning the choices of level of details in the pedagogical agents.

Condition	χ^2	df	p-value
Instructor	18.605	1	< 0.001
Learning Companion	82.046	1	< 0.001

The null hypothesis H_0^4 and H_0^5 are rejected: The instructor ($\chi^2=18.605$, $p<0.001$) and the learning companion ($\chi^2=82.046$, $p<0.001$) versions yield a preference for agents with low levels of detail.

ANALYSIS AND DISCUSSION

Hypothesis 1: Choice of Pedagogical Role for EPAs in ITS

At the end of the study the children were asked to provide rationales for the selection of the EPA with the pedagogical role they most preferred using within *Multipliote*: instructor or learning companion.

The results (H1) show a statistical preference in children's choice of learning companion as EPAs. The reasons behind this choice given by the participants were that they could relate more to the learning companion, and 'trusted' the characters to help them. The instructors were seen as too formal as characters, and most children felt 'judged' by them. This seems to indicate that while children associated with both characters roles, the relation they might have build up with them is of a different nature: learning companions were compared more to 'peer' features, while with the teacher EPA they kept a more formal teacher-student relationship.

Children's preference for a 'peer' feature to a more formal "instructor" one is in line with Acuff & Reiher's (1997) categorization of children by developmental stage with children aged 7-11 years: they prefer to work in pairs or in groups, and this holds here even though the peer is not physically with them, but digitally represented.

Teacher's interviews on the subject revealed that the instructor figures probably represented themselves to the children, along with their style of interaction during a learning session: In their class, when children work in groups they help each other; while when interacting with their instructor, at the teacher's intervention, they only follow the teacher's instructions and advices, but are not looking for extensive and individualized help. Hypothesis can be made that children kept to the working dynamics they held in class, drawing a parallel

with software interaction. From an observation of the differences in children's interaction with different software at home, in a computer science lab outside the classroom, and within classroom settings, the results might have been different if the children had been subjected to the study outside from the classroom.

Hypothesis 2 and 3: Association of Pedagogical Role and Level of Naturalism

The evaluation of H2 and H3 revealed that children conveyed a preference for more naturalistic instructors (humanoid-shaped), and stylized learning companions as EPAs.

Children's preference for more stylized learning companion is coherent with the results of H2 in Haake & Gulz's study (2009). However, unlike their non-significant result when considering the instructor separately, results here appear to be significant: more naturalistic visual appearance is preferred. This could be related to the difference in experimental design: In Haake & Gulz's study, the same characters were used for both conditions, and users chose the visual appearance before selecting the pedagogical role. In this reported study, children chose the characters while distinctively knowing their role. For example, the instructors were clearly identifiable with the presence of a black teacher's hat on their head. In this context, children might have associated the instructors more with their personal reference, their teacher. As seen for hypothesis 1, children seemed to associate the EPA representation with a judgmental character, that would make them work for the answer instead of giving it to them. The characters displayed were not the same in each condition, and were especially designed with this role in mind. For this reason, children may have adopted more naturalistic instructors to be closer to the more formal social role of the instructor, i.e. closer to the reality of the class teacher.

Similarly, the choice of more stylized (smiley-shaped) learning companion could be explained by children's view of the EPA in a more relational social role, like an imaginary friend that needs to be more imaginary than realistic. This corresponds with Reeves and Nass's (1996) Media Equation theory of transference of real world relational strategies into a 'virtual world' – aka computerized learning environment.

The choices made could also relate to the age group of the study participants. Children of this age frequently watch animated movies or read comics, where the simplification of details emphasizes the meaning or semantic association of the images (McCloud, 1993), therefore bringing the user closer, or more associated with the learning companion, and keeping a sense of fantasy, detaching themselves from a totally realistic setting unlike older children.

Hypothesis 4 and 5: Association of Pedagogical Role and Degree of Detail

The evaluation of H4 and H5 revealed that children conveyed a preference for simplified characters for both pedagogical roles. Children's preference for more simplified agents may relate to McCloud's theory (1993) that such design characteristics emphasize social-emotional expression, and facilitate self-identification and immersion into the character of the story.

Results for H3 and H5 concerning the learning companion also correspond with the field observations of Haake & Gulz's study (2009), that when the EPA is associated to a 'friend', they tend to select more simplified and cartoonish characters rather than a more detailed and naturalistic one.

CONCLUSION

The study produced the following results:

- Children aged 7-11 will favour more naturalistic (humanoid-shaped) instructors, and stylized (smiley-shaped) learning companions, both EPAs with a low level of detail in the visual static appearance.
- They prefer learning companions for EPAs pedagogical role, as opposed to a role of instructor/teacher, trusting them more and feeling closer to them.

Some results in this study were similar to Haake & Gulz's study (2009) on older children from a different nationality, such as the preference of more stylized EPAs as learning companion agents (H₂). Other results, however, seemed specifically related to our participants' age group or cultural background (H₄, H₅), and school learning practices (H₁).

Several potential limitations of the study relate to the scope and generality of the results. One limitation is that the children tested the interaction with the agent for a limited amount of time, and within a specific scenario. Although the study is similar in scope to other studies in this area, in advance of making claims about the generality of the results, further studies over a longer time period, and under different conditions of use are necessary. However, these results provided a foundation for the next section concerning with designing an EPA in our OLM, appropriate for children aged 7 to 11 to use, within the educational software.

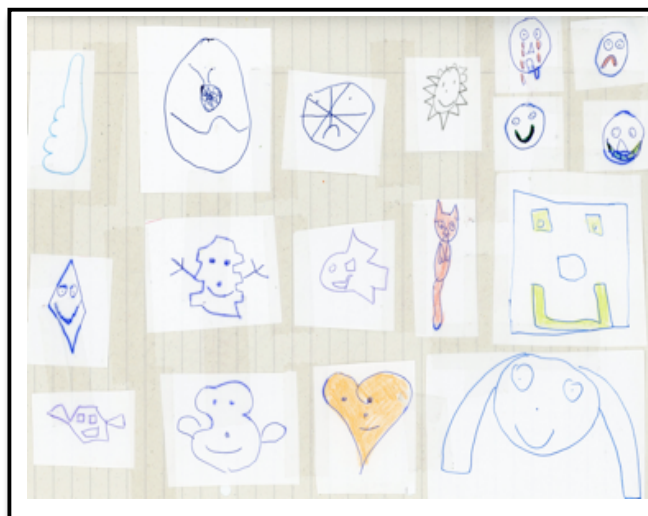
C.2: DESIGNING MULTIPLIO

The participatory-design study involved a French class of 29 children aged 7 to 9 years old. The children had never worked with the evaluators before, nor did they have previous knowledge in learner modelling, or EPA design.

They were presented with an empty array of 10 cells, numbered, and representing each multiplication table from 1 to 10. For each multiplication table, they were asked to think about how well they knew this table, and to find a way to represent that knowledge in the cell through a drawing. They were allowed to use colours, symbols, and text. The activity was not time-limited. There was no restriction on the kinds of iconic symbols they could use for such drawings, and they were not directed in the form the drawings could take.

All participants intuitively chose one particular symbol of their own creation for their drawings, defined from here onwards as “systems”. They altered shape, size, or colour, to represent how well they knew each table.

It appears that the children used 3 distinct metaphors of learning to represent their knowledge of Multiplication tables: the *Traffic-Light*, *Smiley-Faces*, both described in Appendix B, and *Thumbs* metaphors. Each metaphor divides the development of learning for a specific concept into three states. Where the *Traffic-Light* metaphor associates *colours*, and the *Smiley Faces* associates *facial expressions* to each state of learning (not understood, partially understood, understood), the *thumbs* metaphor associates a *position of the thumbs* (up, horizontal, down). Figure C.4 illustrates the kinds of systems generated using a sample of



the icons the children drew:

Figure C.4 Children's representation of their knowledge

Table C.6 represents the association of each child's system representation to one or several metaphors of learning used in European schools:

Table C.6: classification of system representation		
Thumbs	Smiley Faces	Traffic-Lights
1	28	15

One participant used the *Thumbs* metaphor, rarely used in French schools in terms of learning, but commonly used in info-commercials and comics in terms of valence.

All other participants used embodied metaphors with some sort of facial representation that could be assimilated to the *Smiley Faces* metaphor. The shapes of the icons varied from simple geometrical shapes such as circles or squares to human-like characters and pets.

Approximately half of the children (15/29) used colour in their systems, to represent their level of understanding of a particular Times table, thus following the *Traffic-Light* metaphor. The colour semantics used followed the original metaphor, with red for a negative feedback, orange for a medium one, and green for a good understanding. Some of them coloured the totality of the icon/character in the appropriate colour (such as the cat or the heart in Figure C.4), whereas others outlined the emotive expressions on the object/character drawn (such as the faces at the top-right corner of Figure C.4).

Many of the children rated the intensity of their knowledge with a 3-points scale using the following values: bad, medium, good. Moreover, about a third of the children(9/29) used specifically strong emotions for some of the tables, creating a 5-points scale by adding cries and several levels of contentment by varying the width of the smile, or the character, as illustrated in the square of 4 symbols at the top-right corner of Figure C.4.

The results of past studies undertaken with French and English school-children aged 7-11 years old showed that children easily recognize the *Smiley Faces* and *Traffic-Light* metaphors when applied to learning (appendix B). Smiley-o-meters are also widely used in Child-Computer Interaction as a usability tool where children give their opinions, usually following a Likert scale (Read et al, 2002). Furthermore, the results of last section showed that Skill-Meters are the preferred and most likely to help understand the underlying model for children this age group.

From these observations, and the study of EPA's favoured design concerning our users' age group presented in the appendices, a new EPA "Multiplio" was created. Its visual representation aimed at evoking the multiplication symbol, as this mathematical concept is at the core of *Multipliotest*. At first, two versions of the character were designed, as illustrated in Figure C.6:



Figure C.6: The two designs of the multiplication symbol selected

The one on the right represents the computerized version of the multiplication symbol, whereas the one on the left illustrates the hand-script symbol usually used in class. A chi-square analysis of their choice showed a significant preference of children for the computerized symbol ($\chi^2 = 9.9655$, $df = 1$, $p\text{-value} = 0.002$).



After discussion with the teachers on the pedagogical value of choosing this symbol, "Multiplio" took its final form, using the computerized symbol, emphasizing children's recognition of the use on the computer.

In the studies of chapter 5, 6 and 7 concerning *Multipliotest* software, a distinction is made between the system's conception of the LM and users self-identifications, via the use of another icon for the user of the LM: "Moije". The "Moije" character is represented by a "classic" smiley feature: circle for the face, eyebrows, eyes, nose and mouth adjusted to the emotion to portray.

C.3 DESIGN CHARACTERISTICS OF THE EPAs EMBEDDED IN MULTIPLIOTEST AND THE DIVIDINGQUEST.

The design choices concerning all EPAs used in the *DividingQuest* and *Multipliotest* were taken from the results of the literature and the previous study, illustrated in table C.6, and described in the following subsections.

Table C.6: Description of the EPA's design

DividingQuest		Multipliotest
Visual Static Appearance		
Basic Model		
	magical humanoïde creature (witch)	multiplication symbol
Physical Properties	Age: Immortal but looking 15-20; Caucasian ethnicity; clothes, body, and hairstyle related aspects corresponding to the software fantasy theme and used for the inclusion of colours.	No age; presence of a face with eyes, eyebrows, nose and mouth onto the multiplication symbol, which are used in form and colour to represent affect.
Graphical Style	<ul style="list-style-type: none">• dynamic representation;• simplified and low level of details;• stylized and non-realistic;	
Pedagogical role		
<ul style="list-style-type: none">• Expert for navigation and learner-model representation;• Mentor when giving user feedback on his/her progress and facilitating help-seeking behaviour;• Guide and motivator when considering pedagogical choices in interaction;• In overall, moderate authoritative role, learning companion		
Communicative style		
Pedagogue mainly task-oriented, but trying to create an affective relation with the user: low level of argumentativeness; high degree of assertiveness; more descriptive than narrative; extrovert; and quite connected to the user.		

VISUAL STATIC APPEARANCE

The basic models were chosen to be integrated within the fantasy concept of the educational application: a whole storytelling has been created around the learning of divisions in the *DividingQuest* (Girard, 2006), facilitating a help feature to be symbolized by a prototypical basic model of a human as an immortal magical creature, very much like the avatars in most popular and engaging video games of Fantasy (FinalFantasy™, World of Warcraft™, Dofus™, ...). While this character could be mistaken for naturalistic due to its

human form, children in participatory-design sessions associated it more to a cartoonish figure, emphasized by the simplified level of details, really embedded in the fantasy and therefore far from belonging to the real world. This complies with the results of the study presented in the previous section, with children favouring more stylized characters with a human-shaped form, for an EPA that holds some aspects of a teaching pedagogical agent. Furthermore, when asked separately to choose whether it represented a companion or more of an instructor figure, just on visual representation, they selected the companion at 100% (20 out of 20).

On the contrary, *Multipliotest* software contains simplified interface design, dealing only with multiplications without the presence of a background storytelling. During the participatory-design sessions when developing *Multipliotest*, all children involved (29 aged 7-9 years old) drew non-human smiley-like characters, embedded within forms of animals or mathematical symbols. This led to the creation of ‘*Multiplio*’, as illustrated in Table IV.6, this cartoonish, smiley-like, and simple in design character, in the form of a multiplication symbol, with a face that can smile and cry like a child.

French and English primary school teachers do not follow the same curriculum in mathematics apprenticeship, or pedagogical methods. For this reason, two OLM applications were developed, each designed including the cultural and pedagogical characteristics for use by French and English children aged 7 to 11 years. The choice in the (non)-human representation of the characters was decided in relation to the storytelling defined during participatory-design sessions to build the applications. Within the *DividingQuest*, the “world” is defined as a fantasy game setting, with sceneries embedded in the conceptual maps, which led to the integration of fantasy characters that were embodied rather than just smiley-like shapes. On the contrary, the *Multipliotest*, targeting younger children, is represented by an interface sticking to the multiplication metaphor, with no real storytelling behind it other than a rewarding scheme in the drill-and-practice component. For this reason, and from the results of the study presented in the previous study, the character chosen was designed as a smiley-shaped multiplication sign, to represent what children defined as “the multiplication master” character, *Multiplio*.

Physical properties:

It was decided to scope this PhD’s work on “a digital character”, without any consideration of gender. Gender was therefore not included as factor in the experiments and no assumptions will be made as to its impact on the experimental results given from the

DividingQuest. There is a unique character, which is more focused on the magical aspect. It was decided of female gender to emphasize the light ‘expert/instructor’ aspect of the learning companion character, as only female characters were considered by children for this pedagogical role of EPAs in our study (cf. section C.1). The results of the *DividingQuest* will be analyzed for this particular character, and the changes of such results when using characters of other gender, ethnicity, age, or other changes in visual static appearance will not be investigated here. For *Multiplotest*, the character designed is asexual, and does not “speak”, so there should not be any gender issues involved. All elements of the EPA’s physical properties are used to embed affect into the EPA’s design through the use of colours and facial/bodily expressions, respectively the Traffic-Light and Smiley Faces metaphors of learning according to the studies design.

Graphical style:

Dynamic visual animations were chosen over static visual pictures as it generally reinforce the experience of anthropomorphism by mimicking schemas for human behavioural and communicative strategies (Haake & Gulz, 2009), therefore creating a character culturally closer to the children.

Designers’ choices regarding the physical properties of EPAs always have social, cultural, psychological, and affective impacts on users and how they manipulate the interfaces (Haake & Gulz, 2009). The design aimed at limiting comparisons between characters on their form, fashion (hair or clothing), or body morphing, to avoid users’ reactions to attractiveness, personality, or gender stereotypes. Our unique human-like EPA has been designed with the following characteristics: average proportioned in terms of body mass and height, fashion clothed (according to children’s representation of what a witch’s clothes should be like and their fashion trend, in order to avoid fashion issues and problems of identification as described in (Isbister, 2006). Clothing as a concept is not represented in the *Multiplio* character, which graphical representation only vary in the form and colour of its facial attributes.

Both characters are simplified with a low level of detail to increase the distinction of facial expression, at the centre of our investigations. This reduction of details has been found to sometimes support a more rapid and accurate processing and interpretation of the character’s features (Cook 1979, Isbister 2006), as well as facilitating the user’s self-identification with the character, and therefore increasing his or her involvement in what the character says or does (Gulz & Haake 2006, McCloud 1993).

No definite answer on the impacts of ‘realism’ in characters on user responses to EPAs and the interfaces they are embedded in (Haake & Gulz, 2009), currently exists. Regardless, the design choice of our characters being ‘non-realistic’, simplified, stylized graphical representations with a medium level of visual details is based on McCloud’s approach (1993). It states that users will be more likely involved and able to project themselves into a visually simplified character, as the stylized character will invite the user to elaborate on its story and design with his or her own personal and subjective experience (Haake & Gulz, 2006). The results of the previous study tend to agree with that statement, with children choosing the “x” agent as learning companion because of its correlation with the mathematical content of the software: the symbol, representing the multiplication symbol when children work on multiplications on paper, was identified as a “multiplication master” by the participants. The identification with the character added another level to the meaning of the agent, reinforcing the underlying pedagogical role. According to McCloud’s theoretical framework (1993), cartoonish and simplified characters can help amplify the meaning of an image, for example the meaning of facial expressions that can influence socio-emotional communication, fundamental element in companionship.

PEDAGOGICAL ROLE

Our characters have both the physical and pedagogical descriptions of an expert, mentor and motivator as described in Baylor and Kim (2005), and were based on the results of their described experiment, as well as our study’s results on children’s preferences in EPAs (cf. section 4.2.1):

- Expert : increasing information acquisition processes by giving an expert role to the navigation help, and visualization of the learner model;
- Motivator role on the choices made during a learning session to increase children’s self-efficacy;
- Mentor role when giving answers or offering mathematical help in order to improve learning and motivation.

When considering the encouragement given to the users, the characters therefore are both informative and encouraging, with a variety of emotional representations with a mixture of dietic gestures and emotional expressions, and highly-animated gestures.

COMMUNICATIVE STYLE

Finally, the communicative style was inspired from Bickmore's task-and-relation-orientated and task-orientated categories of EPAs (Bickmore, 2003). In the first category, characters are more focused on task and the learning environment to be close to the Baylor and Kim's (2005) Mentor agent: the user and the agent are working closely together for the Mentor to demonstrate competence while simultaneously developing an affective/social relationship to motivate the learner. However unlike the description of task-and-relation-orientated in Bickmore (2003), the EPAs here do not share personal information with the children, but aim at creating an affective link with them for the time of the learning session only. The *DividingQuest's* witch and Multiplio are therefore considered as task-and-relation-orientated EPAs, with a relation-oriented strategy limited to creating an affective link with the learner.

The use of stylized representation for task-and-relation agents EPAs also corresponds to what English 12-15 years old users would instinctively choose as EPA helpers in an educational software (Haake & Gulz, 2009), is revealed similar to our initial findings on user preferences with children aged 7-11 favouring learning companions because of their relational potential, and finding teacher-like figures too close to the task.

The results and decisions made in this section helped the creation of two EPAs, one for each piece of pedagogical software used in this PhD work. The aspects of visual appearance, including the level of details and naturalism of the characters, as well as the communicative skills and pedagogical roles were investigated. The results from chapter 3 produced a model of EPA's emotive responses to software interaction to integrate within the OLM environments in terms of pedagogical feedback to user activities. The last component to be integrated in the educational software is concerned with how OLM components can be designed so as for children to understand their underlying concept, and be more willing to interact with said components. In the next section, some aspects of learner-modelling for children's use will be investigated.

APPENDIX D: ANALYZING CHILDREN'S INTERACTION STYLES IN EDUCATIONAL SOFTWARE : POINT-AND-CLICK VERSUS POINT-AND-SELECT.

In the context of developing and evaluating software for child users, studies have been investigating different interaction techniques used by children, and the user's accuracy in using the systems within such interaction styles. This study is interested in investigating the usability of two commonly used interaction styles: point-and-click, and point-and-select, to help in the development of the applications used in this thesis for the empirical studies.

METHOD

The experiment examined children using versions of Multiplotest with either a point-and-click interaction style or a point-and-select interaction style. The goal of this experiment was to determine the impact of these in-teraction styles on children's achievement, motivation and preference, speed of answer, and accuracy of answer. Each child used both versions of the software, with a 24 hours gap.

PARTICIPANTS

24 children from the class (aged 7 to 9 years) of level CE2 participated in the study by playing two versions of the software. Each child used both versions of the software, with a 24 hours gap, the activities in between sessions excluding any mathematical content. The order of experiencing the different IS was counterbalanced.

TWO INTERACTION STYLES FOR ONE TASK MODEL

The comparison of interaction styles within Multiplotest was related to one key task in the game: giving the result of a multiplication. Children from school level CE2 learning multiplications for the first time still break down numbers into columns, describing them as a succession of digits : units, tens, hundreds, ... For this reason, the software asks the children to give an answer by selecting the number corresponding to each of the digits of the answer.

Figure 1 shows the task model (using the K-MADe notation [18]) corresponding to this task : « Enter the result».

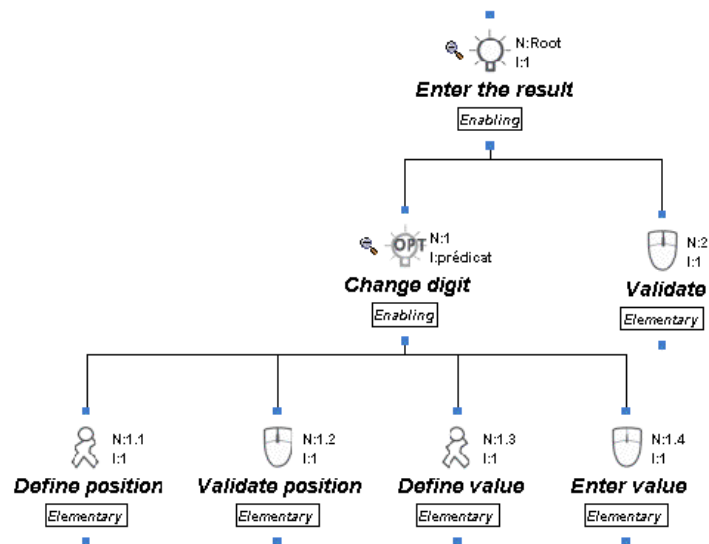


Figure D.1 : The K-MAD task model for the activity of “Entering the result of a multiplication”

When giving the result, the child has to enter each digit separately before validating his/her choice (validate task). The validation can be done anytime after changing one digit. The child is allowed to pass one question entirely and leave the digits to their initial value 000 (Optional characteristic of the change digit task. To enter the value of one digit (change digit task), s/he must first define the position of the digit s/he wishes to modify (units, tens, or hundreds) (the define position user-task), validate the position (the validate position interactive-task), define the value by choosing the number corresponding to this digit (integer between 0 and 9) (the define value user-task), and then enter this value (the enter value interactive-task).. The child can change the value of each digit several times before validating (iterative characteristic of the change digit task), so as to be sure of his/her answer. The validation ends the enter-the-result activity and leads the user to the solution of the multiplication, and his/her feedback on the answer given.

The differences produced by the two interaction styles, point-and-click (Figure 2a) and point-and-select (Figure 2b) are present in two interaction tasks : Validate position, and Enter value.

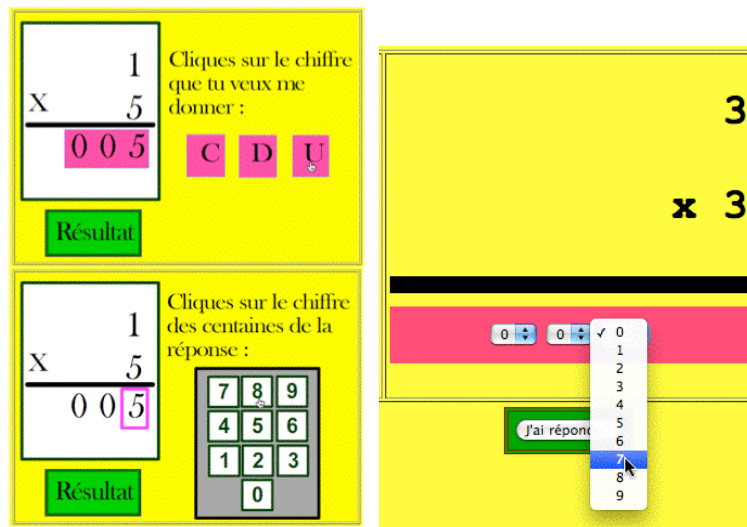


Figure D.2 : Point-and-click activation (left) and point-and-select (right)

In the version with point-and-click (Figure 2a), the child moves the mouse over the column s/he wishes to modify (units column in Figure 2a, top screenshot), and then clicks on the square (Validate position task). A new screen appears (Figure 2a, bottom screenshot), and the user can then choose the integer (from 0 to 9) answer to the question, and then enter it by using the calculator : s/he moves over the number, and then clicks on the square (Enter value task).

The Validate position task within the version with point-and-select (Figure 2b) is quite similar to the previous one, as the child moves the mouse over the column chosen and clicks on the corresponding tab. A list of integers (from 0 to 9) appears on the same screen, and the child can then choose and select the integer answer to the question by moving over the number and clicking once (Enter value task).

Both versions of the software therefore require the same number of clicks with the mouse (two for each digit modification and one for the validation), and changes within the interface. Once the answer has been given, the child clicks on the validation button (green in the interface as shown in Figure 2a and 2b), validating the answer.

HARDWARE AND SOFTWARE

All three computers used were Macintosh Mac Book portable computers (Colour 13-inch screens), with a ‘two-buttons’ mouse. Most of the interactions with the software were undertaken using a mouse, with the exception of the login procedure, which required the use of the keyboard. While all children had a two-button mouse, the three mice used were not identical size-wise. However, Hourcade et al’s (2007) investigation of the impact of mouse

size on accuracy and efficiency of pointing tasks did not reveal any significant difference. It is therefore unlikely that most of our participants would perform differently given different mouse sizes.

PROCEDURE

The children performed individually with software, placed in the experiment room three by three for 20 minutes, supervised by a researcher. The material was placed in such a way that children did not see the others' work, so as to not get distracted by it. A short introduction to the game and the experiment was undertaken prior to the experiment in front of the whole class. Once logged-in, the children were free to use the software by choosing the multiplication tables and levels to try and master. The researcher was monitoring the learning session, and helped the children when the questions were related neither to multiplications nor interaction. During the session, the computer recorded the time taken to answer questions, as well as the number of interaction errors whilst giving an answer. Once the 20 minutes were over, they went back to class for another group to take their place. The next day, the same groups were asked to come again for another session, each child using the IS they did not use the first time. At the end of the experiment, the children were asked to rank their preference for IS using a 5-points smiley-o-meter (Read et al 2002). For each IS, the child had to define how much s/he found it useful for giving answers.

TWO INTERACTION STYLES FOR ONE TASK

Multiplotest, developed in collaboration with French teachers and primary school children, contains three activities in order to learn how to multiply and to revise multiplication tables. In this study, children could interact with and change their user model using skill-meters, revise their multiplication tables by requesting a separate help screen, and choose the test they want to do (restricted to level 1 – one multiplication table - and 2 – all multiplication tables from 1 to 10 - because of their limited knowledge of multiplication tables at the time of the experiment.) The comparison of interaction styles within *Multiplotest* was related to one key task in the game: giving the result of a multiplication. Children from school level CE2 learning multiplications for the first time still break down numbers into columns, describing them as a succession of digits: units (U), tens (T), hundreds (H), ... For this reason, the software asks the children to give an answer by selecting the number corresponding to each of the digits of the answer. The differences produced by the two interaction styles, lies in the

validation of the digit to modify, and the selection of the actual number given for said digit. In the point-and-click version, the children first click on the icon selecting the column considered, before moving the mouse over the calculator to select the digit corresponding to the column. In the point-and-select IS, the children also click on the column to modify, opening a sliding bar for this column, and then slide the highlighted number vertically until they reach the digit desired.

Both versions of the software require the same number of clicks with the mouse (two for each digit modification and one for the validation), and changes within the interface. The widgets used for both IS were equal in size, and always fixed in position on the screen. However, while for point-and-select, the selection occurred where the digits columns were, for the point-and-click interaction, the calculator was on the right part of the screen, separated from the answer area.

EXPERIMENTAL DESIGN

In this experiment, one independent variable was manipulated, the mouse interaction style (two levels: point-and-click, and point-and-select). The questions to be answered in each version of *Multipliotest* were identical, except for the IS used to answer. Each child used both versions of the software, one group beginning by the point-and-click version, and the other by the point-and-select version. The children were randomly assigned to one of the two groups, but their mathematical abilities in regards to multiplications were taken into account for the two groups to be balanced.

Four dependant variables were measured: children's learning achievement; preference; overall movement time; and interaction errors. The results concerning children's achievements, overall completion time, and interaction errors were analyzed using a two-sample correlated T-Test with 23 degrees of freedom. The null hypothesis states that there is no difference in interaction styles concerning the dependant variable. The alternate hypothesis is that the interaction style impacts on the results of the dependant variable.

RESULTS

ACHIEVEMENTS

Table D.1 shows the number and percentage of children answering none of the questions successfully, and Table D.2 the number of questions answered correctly.

Table D.1: Number of children who were unable to solve any question

Condition	Zero correct	%
Point-and-click	1	4%
Point-and-select	8	33%

Of all the children participating in the study, 33% of them were not able to answer any question correctly when using the point-and-select version, and only 4% in the point-and-click version of the game. As there was no difference in mathematical difficulty of the questions to be answered for both IS, it seems that the problems of use of the IS had some impact on their ability to answer questions: while they knew the answer in most cases, they were not able to enter it, and preferred to pass the question after struggling too long.

Table D.2: Number of questions correctly answered

Condition	Mean	S.D.
Point-and-click	2.18	1.80
Point-and-select	1.416	1.61

While all children answered between 0 and 6 questions correctly, when using the point-and-click interaction style, they were able to perform significantly better than in the point-and-select condition, $t(23)=8.577$, $p=0.001$.

OVERALL COMPLETION TIMES

The average completion times for the two interaction styles are shown in Table D.3. Only questions where no error occurred were included in the means. When using the point-and-select IS, children were found to take a significantly higher number of seconds to perform the task, $t(23)=20.783$, $p=0.001$.

Table D.3: Overall movement time

Condition	Mean	S.D.
Point-and-click	6.58	1.81
Point-and-select	27.2	3.27

The average number of errors for each interaction style is shown in Table D.4. The analysis revealed that the point-and-select IS led to children committing a significantly higher number of interaction errors, $t(23)=11.281$, $p=0.001$.

Table D.4: Number of interaction errors

Condition	Mean	S.D.
Point-and-click	0.66	1.62
Point-and-select	6.66	2.76

PREFERENCES

Children's preference of interaction style can be grouped into three nominal categories: prefer point-and-click, no preference, and prefer point-and-select. 87.5% of the children preferred the point-and-click IS (21/24), 8.3% had no preference (2/24), while only 0.4% preferred point-and-select (1/24) as it required less mouse-movement across the screen.

DISCUSSION

IMPACT ON LEARNING

The results from Table IV.8 demonstrate that a point-and-click interaction style, used in an interactive learning environment, can be more effective in terms of performance than a point-and-select IS. While one could argue that the knowledge of multiplications might have evolved during the 24 hours gap, and influenced by the first use of the software, the data was counterbalanced in that aspect as half of the children began with point-and-click, while the others began with the point-and-select version. No effect was found in the order in which the versions were used. However, as this experiment manipulated the styles within a rich, complex environment, some other factors might have interfered with the results (such as the choice of activities to perform).

OVERALL COMPLETION TIMES AND INTERACTION ERRORS

The results of the experiment highlight that point-and-click is a more effective mouse IS than point-and-select, both in terms of speed and accuracy. The completion time results included only the questions in which no interaction errors occurred. While these results demonstrated that the point-and-click interaction style was faster than the point-and-select interaction style, the children made errors in approximately 26% of the questions. By including those questions in which errors occurred, differences between the movement times for point-and-click and point-and-select increased significantly ($t(23)=47.359$, $p < 0.001$).

Instead of an average difference of 36ms between the two interaction methods, the average difference increased approximated 230% with the point-and-select interaction style becoming on average 120ms slower than the point-and-click IS.

The errors due to the interaction were mainly due to the selection within the point-and-select version. Children indeed had difficulties moving the mouse down the list of numbers to accurately select the number they wished to choose. Observations from the study showed that the level of frustration increased along with the use of this version of the interface. This resulted in children missing questions rather than trying to select the correct number. The number of errors in this version may also be linked to the representation children had of the task. In the point-and-click version, they were familiar with the metaphor of the calculator, used to visualize all numbers at once and select one among them. In the select version however, they had to “uncover” the number, and select it in a more precise manner. Furthermore, the point-and-select IS can be related to the drag-and-drop IS as it requires persistent pressing of the mouse in order to go on the digit chosen, and then let go of the mouse button. In the point-and-click IS, children had more flexibility of movement between selecting the column to modify, and the digit to enter.

PREFERENCES

Children who used the point-and-select version first were more decisive in their ranking of the interaction style, and selected the extremes in both scales (extreme positive for point-and-click, and extreme negative for point-and-select). Many children who preferred the point-and-click interaction style explicitly stated that they found the point-and-click easier, as when they had problems guiding the mouse to the selection area, they could try again more easily with this IS. They also complained that the point-and-select interaction style “gave them wrong answers”.

CONCLUSION

This study results investigate the usability of the point-and-click and point-and-select interaction styles. The investigation took place within a field study investigating children’s natural interaction whilst playing two versions of an OLM learning environment, *Multiplotest*, where each version only supported one interaction style. The results highlighted that point-and-click is a more effective mouse interaction style than point-and-select for

children aged 7 to 9 years, in terms of achievement, interaction error, speed and accuracy of answers, and children preference.

While the data analyzed only represents one snapshot of effects of user interface style within a learning environment, a design choice was made to follow such results as well as Inkpen's (2001) on the drag-and-drop IS, and to integrate point-and-click as the only IS present in all software to be experimented in this thesis.

APPENDIX E: HOW WELL DO PEOPLE UNDERSTAND AND CLASSIFY EMOTIONS?

Research in affective computing has been increasingly interested in the potential of using different aspects of affect in the design of educational technologies. Studies in the technology-enhanced community have used different models of emotion, from Ekman's six universal basic emotions (1972), or Ortony, Clore, and Collins's cognitive model (1988), to sets of emotions considered relevant to learning (Kort et al, 2001). However, research in child psychology and development has proven that the understanding of emotional concepts and representations grow with age, developed through maturation and social contacts with others (Izard et al, 1986, Stein et al, 2000). For this reason, any model of emotions used in pedagogical software should take into account the emotional literacy of the targeted users, in this case children aged seven to eleven years. In this section, a study is presented comparing children's and adult's understanding of emotions represented by digital characters. The system used to test the understanding was originally designed for international adult users. Results from the investigations as to the differences in recognition and understanding between adults and children will then be used in the definition of emotional representations of EPAs to be integrated in OLM systems, as well as the creation of a self-report tool to capture and represent children's emotional state, presented in chapter 2.

AIM OF THE EVALUATION

The aim of this study was to identify any differences between the three children's age groups, and between the children and adults, in recognising graphical representations of emotions in the form of digital characters.

RESEARCH QUESTIONS

The research questions to be pursued are:

- Which emotions are better recognized by the different age groups?
- Which emotions are confused the most, and by whom?
- Do the different age groups confuse these emotions with the same concepts/graphical representations?

THE PREMOTM SYSTEM

The “PrEmoTM system” (P.M.A Desmet, Hekkert, & Jacobs, 2000) is a self-reporting method used for usability studies concerning the design of products with emotional value. It measures 14 emotions often elicited by product design, each portrayed by an interface character using dynamic facial, bodily, and vocal expressions. Out of the 14 emotions, 7 are of positive valence, and the other half of negative valence, when considering the dimension of “pleasantness”, as illustrated in Table E.1.

Table E.1: The emotions of the PrEmoTM system.

Pleasant emotions	Unpleasant emotions
Fascination	Boredom
Satisfaction	disappointment
Admiration	dissatisfaction
Amusement	unpleasant surprise
Inspiration	Disgust
pleasant surprise	Contempt
Desire	Indignation

PrEmoTM aims at being used cross-culturally and does not ask users to verbalize their emotions. It has previously been found that the correlation between emotions and their graphical representation for adults when doing usability studies has an error rate of less than 5% (P.M.A. Desmet, et al., 2001). Any emotional affect instrument should be usable by both children and adults, from different nationalities, and should not create difficulties of use for children of low-literacy level.

PARTICIPANTS

The study involved 39 French children (25 aged 6 to 8 yrs, and 14 aged 9 to 13 yrs), and 25 international adults.

DESIGN OF THE STUDY

The study included one between-subjects, and one within-subjects, independent variables:

- One between subject variable *age group* with two levels when comparing the 25 6-8 yrs olds children and the 25 adults, and four levels when comparing the children’s age groups.
- One within-subject variable with fourteen levels, which comprises the emotions to classify: *desire, pleasant surprise, inspiration, amusement, admiration, fascination, satisfaction, indignation, contempt, disgust, unpleasant surprise, dissatisfaction, disappointment* and *boredom*.

There is only one dependent variable, representing whether a card has been correctly classified, in the form of a label corresponding to the verbal form of emotion chosen. For example, Participant 23 placed the card representing the emotion '*desire*' on the written label '*fascination*'. It will therefore be coded as the variable 14 ('*desire*') holding the value 8 ('*fascination*') for the 23rd user case. However, when investigating which emotions are better recognized, the dependant variable is binary (0=incorrectly recognized, 1=accurately recognized).

PROCEDURE

Each participant was presented with a pile of 14 cards, randomly mixed. Each card, taken from the PrEmoTM system, represents one of the 14 emotions studied (presented in *section 3.2.5*).

The participants were requested to study each individual card, and to place it on the emotion they thought it best represented. The concepts behind the graphically represented emotions, written on white clouds, were randomly positioned in a different order for each participant, on a board in front of the participants.

There was no time limit for them to complete the task. The participants were allowed to take back a card already placed on the board to place it somewhere else if they changed their minds. Also, several cards could be placed on one emotion, and it was possible for emotions not to have any associated cards.

RESULTS

RECOGNIZING EMOTIONS ACCORDING TO AGE GROUPS:

The research question concerns whether people recognize and classify emotions differently as they develop and mature. Table E.2 illustrates the percentages of correctly placed cards for the different age groups.

Table E.2: Classification of correctly-placed cards by age group (percentage)

Emotion	Adults		6-8		9-13		Total	
	%	SD	%	SD	%	SD	%	SD
1. boredom	84	0.494	68	0.504	88	0.512	77	0.500
2. disappointment	72	0.505	48	0.504	88	0.512	62	0.502
3. dissatisfaction	36	0.478	44	0.492	38	0.403	43	0.481
4. unpleasant surprise	88	0.443	48	0.483	25	0.341	66	0.490
5. disgust	80	0.484	72	0.492	75	0.512	66	0.502
6. contempt	48	0.534	48	0.504	63	0.512	44	0.495
7. indignation	52	0.490	52	0.496	38	0.477	53	0.484
8. fascination	68	0.504	16	0.455	38	0.403	41	0.457
9. satisfaction	32	0.404	40	0.414	25	0.341	32	0.388
10. admiration	12	0.303	20	0.446	0	0	15	0.424
11. amusement	40	0.453	44	0.471	75	0.516	44	0.484
12. inspiration	64	0.490	20	0.436	25	0.447	41	0.457
13. pleasant surprise	32	0.418	28	0.436	50	0.478	41	0.481
14. desire	40	0.418	36	0.477	50	0.5	39	0.477

Three negative emotions are recognized the best when considering percentages for all participants: *boredom* (77% of correctly-placed in total), *unpleasant surprise* (66% in total), and *disgust* (66% in total), closely followed by *disappointment* (62% in total). On the contrary, *dissatisfaction* and the positive emotions *satisfaction*, *admiration*, *amusement*, *pleasant surprise*, *desire* are poorly recognized by all the participants, with a percentage of correctly placed cards for all participants ranging from 15% to 44%. While the users seem to be able to recognize the valence of the emotions, the instrument does not seem to enable them to distinguish and individually recognize the emotions within the positive valence group.

The results of a chi-square test on the total number of emotions recognized by the participants indicates that the distribution of the adults and middle-aged children is significantly different from what would have been expected by chance alone ($\chi^2=20.767$, $p=0.008$). Inspection of the data revealed that adult participants recognize approximately 7 or more emotions, while most of the children were classifying correctly less than 5 emotions, with the exception of 2 children (12.5%) who recognized 9 emotions out of the 14. The two children performing best were the eldest of the 9-13 yrs group, a few months from their 14th birthday.

However, a chi-square test performed on the three children' groups and considering the total number of emotions recognized failed to reveal any statistical difference for the child population alone.

Chi-square tests were performed on each emotion classified by comparing the 6-8 yrs old children and adult population, and then the two children age groups. The results were significantly different for the following emotions, presented in table E.3:

Table E.3: Chi-square rejecting the null hypothesis when considering emotions separately.

Emotion	6-8 yrs old vs. Adults		6-8 yrs vs. 9-13 yrs	
Unpleasant surprise	$\chi^2=9.191$	$p=0.002$	N.S.	N.S.
Disgust	N.S.	N.S.	$\chi^2=14.055$	$p=0.001$
Contempt	N.S.	N.S.	$\chi^2=7.565$	$p=0.023$
Fascination	$\chi^2=13.875$	$p<0.001$	N.S.	N.S.
Inspiration	$\chi^2=9.934$	$p=0.002$	N.S.	N.S.
Pleasant surprise	N.S.	N.S.	$\chi^2=12.688$	$p=0.002$

When looking at the recognition rate of *unpleasant surprise*, the results are contradictory and show a significant difference between the adults and the 6 to 8 yrs populations: whilst being the best recognized emotion for adults (88%), it was only recognized by 48% of the 6-8 yrs children group. Similarly, *fascination* and *inspiration* produced contradictory results and show a significant difference in recognition of the emotion between the adults and 6-8 yrs old children population with a percentage of recognition ranging respectively from 68% and 64% for adults, to 40% and 20% for children. However, the results failed to reveal any difference between the two children's age-groups, with percentage rates ranging from 16% to 38%.

The emotions *contempt* and *disgust* were well recognized by the children, with *disgust* was the best recognized emotion for the 6-8 yrs (72%), and well recognized by the eldest group (75%), *contempt* well recognized by both groups (48% and 63%, respectively).

In considering the percentages for all participants, it can be seen that the positive emotions (15% to 44%) were less well recognized than the negative ones (43% to 77%). No difference in population was found between any of the user groups concerning the negative set of emotions (cf. section 3.1.3 in the Appendixes). However, the results of the recognition of positive emotions are significantly different between the child and adult population ($\chi^2=21.600$, $p=0.001$), and between the children groups ($\chi^2=24.329$, $p=0.007$).

There is a tendency for adult participants to recognize 5 out of the 7 positive emotions. The spread of the 6-8 yrs old is however widely distributed with most of the children recognizing less than 4 emotions, while the 9-13 yrs old usually recognized 4 emotions or

more. However, the percentages of recognition and the size of the user samples are too small to infer a developmental difference in positive emotion recognition by children and adults. The difference in percentages could be explained by the instrument itself, as it was designed for adult users.

When comparing the results of the adults with the children groups, we can see that the adults performed better on all positive emotions. However, while adults also performed slightly better on the negative emotions, the difference in recognition was not significant.

CONFUSING EMOTIONS:

The proportion of cards misplaced for each of the 14 emotions was quantified. Table E.4 indicates which emotions were most confused with which others, by each age group. The numbers underlined represent the emotions most chosen in place of the actual emotion represented (and chosen by at least 20% of the participants).

Table E.4: Classification of which emotions were confused for which by age group (numbers of the emotions by at least 10% of the participants)

Emotion	Adults	6-8	9-14
1. boredom	-	<u>2,5</u> , 8, 9, 10	8
2. disappointment	1, 3, <u>4</u> , <u>5</u> , 6, <u>7</u>	<u>1</u> , 5, <u>8</u>	9
3. dissatisfaction	4, 5, 6, 7	5, <u>6</u> , <u>7</u> , 10	<u>5</u> , <u>6</u> , 7
4. unpleasant surprise	5, 6, <u>7</u>	<u>2</u> , 5, <u>7</u> , 10	2, 6, <u>7</u>
5. disgust	7	1, <u>4</u> , 6	4
6. contempt	<u>3</u> , 8	1, <u>3</u> , 7	1, <u>3</u>
7. indignation	2, <u>3</u> , <u>6</u>	3, <u>4</u> , 6,	<u>3</u> , <u>4</u>
8. fascination	9, <u>11</u> , 13	<u>9</u> , 10, <u>12</u> , <u>13</u>	<u>9</u> , <u>10</u> , <u>12</u>
9. satisfaction	8, 9, 11, 13	<u>8</u> , 11, <u>12</u> , <u>13</u> , 14	9, 13, <u>14</u>
10. admiration	13, <u>14</u>	9, 14	14
11. amusement	<u>10</u> , <u>14</u>	9, 10, 13, <u>14</u>	10, 14
12. inspiration	9, 13, 14	<u>10</u> , <u>11</u> , 13	<u>10</u> , 11
13. pleasant surprise	9, <u>10</u> , 14	<u>8</u> , <u>10</u> , 11, 12	9, 10
14. desire	9	<u>8</u> , <u>9</u> , 10, 12, 13,	<u>12</u>

The most confused emotion, *dissatisfaction*, is confused with *disappointment* (and vice versa), *contempt* (and vice versa), and also with *disgust*, by all groups. The three most inaccurately recognised emotions, *fascination*, *satisfaction*, and *admiration* are confused with positive emotions. Whilst *amusement* is confused by all groups with *admiration*, this emotion is mainly confused with *desire*, again for all participants.

All children groups confused some of the 4 first negative emotions in Table E.4 (*boredom*, *disappointment*, *dissatisfaction*, *unpleasant surprise*) with three positive emotions, *fascination*, *satisfaction*, and *admiration*, by 6 to 8 year olds. Similarly, there is a presence of

the emotion *indignation*, recognized by 53% of the participants, in the three emotions adults and children confused it with (*disappointment*, *dissatisfaction*, *unpleasant surprise*).

DISCUSSION AND CONCLUSION

The study aimed at investigating developmental differences between adults and 2 children age-groups in emotions recognition of the PrEmo system digital characters, simulating 14 different affective states.

While most negative emotions were generally well recognized by all participants, the positive emotions were identified very poorly by all participants, and confused with emotions from the same valence. Furthermore, the four best recognized negative emotions were also the most confused in the negative polarity of the emotions. In average, children seemed to be able to recognize far less emotions than the adults. The results seem particularly contradictory between the adult and child populations for three emotions: *unpleasant surprise*, *fascination*, and *inspiration*. While adults rated a high recognition score (over 60% in all cases), between 50 and 80% of the children failed to identify them. It can be concluded that while the PrEmo system allows users to recognize with accuracy the valence of the emotional representations, the emotions within each polarity are easily confused by young children with one another. It would be interesting to know whether the misconceptions detected lie in the particular graphical representation of the emotions, or in the understanding of the underlying concepts in these positive and negative emotions.

Interviews of the child participants revealed that some children aged 6 to 8 focussed entirely on the facial expressions, while others took only bodily postures into consideration. Some of the youngest participants did not associate well with the characters because of the lack of reality in the body proportion (“the head is too big, it’s a clown...”, and a clown can never be sad or cry, right?). The emotions *captivated* and *disgust*, in particular, were found hard to recognize because of the shape of the mouth: the tongue in *captivated* was not always recognized as such, and the expression of the mouth in *disgust* was very hard to read for some, which hindered the understanding of the expression for some of the youngest participants. On the contrary, some of the body postures helped greatly in the recognition of the emotions (for example, *unpleasant surprise* or *dissatisfaction*). The lack of homogeneity between the body/facial parts used in the design of the affective states was also found troublesome to some child and adult participants: when new parts were detected by the participant, s/he reviewed the whole pack of expressions, inspecting them more closely to find such aspects in the

pictures, which distracted most young participants from the task at hand. At last, the number of emotions to identify was found too high by some of the youngest participants. Some of them got bored with the activity, getting more and more distracted.

In conclusion, it would seem that the visual cues and affective facial/bodily expressions offered by the PrEmo system are not sufficient to help children recognize the 14 emotions represented, especially with regards to the positive emotions. In this chapter, additional information will be given to distinguish the various emotions employed, by using a colour-code, described in the next section. The research hypothesis behind this colour-code is that a combination of coloured and affective visual cues, designed with and for children, will help in the understanding of the affective aspects of the emotional states considered.

APPENDIX F: DESIGNING EMOTION SETS FOR PEDAGOGICAL SOFTWARE WITH CHILDREN AND TEACHERS

Designers in the CCI field advocates that involving children in different phases of the design process is vital to ensure that most of the requirements needed are included (Mazzone & Read, 2005). The results of the study described in section 3.2 showed that children and adults do not recognize, or understand, affective states at the same level. Therefore, when designing affective educational products, using children from the product's targeted age-group should be beneficial to assure a conformity in the desired and actual product use. However, when considering using affect to represent educational feedback, children alone are not enough to design a product of good pedagogical value. The inclusion of teachers working with children of the appropriate age group should help in the definition of an emotional set corresponding to the pedagogical goals pursued. In this section, a set of emotion is defined using children as design informants. It aims at creating a set of emotions children might feel during a learning session on technology-enhanced software.

DEFINING EMOTIONS FOR SORÉMO: CAPTURING THE USER'S EMOTIONAL STATE

This section presents an overview of the design process of the instrument with the selection a set of emotions to form the emotional state of a player.

ITERATIVE, USER-CENTRED, PARTICIPATORY-DESIGN APPROACH

An iterative user-centred participatory design approach was chosen to construct *Sorémo*. Each iteration of the design of the graphical representations of emotions to be included in *Sorémo* followed the same procedure, the studies only varied in number, age, and nationality of participants; and the emotions studied. The design process involved children from French and English primary schools, as well as international adults, as testers and for some of them as design partners.

The PrEmoTM system uses 14 emotions. Children have limited cognitive capacity that causes them difficulties in focusing on this many concepts. Even assuming it was possible, it would inflict long and complex memory processing, distracting them from any other task at hand (Bull et al, 2005). Taking account of the results from the study presented in appendix E,

and the need to overcome problems with the number of emotions chosen for representation, the number of emotions to be used within the new method was reduced to nine: four positive, four negative and one neutral emotion. In reducing the number of emotional response representations, it is intended that the new self-reporting method will be less disruptive and distracting for learning processes.

DESIGN AND SET-UP OF THE STUDIES

Research questions

The following research questions were of interest:

- Are the emotions well recognized?
- Which emotions are not well recognized?
- Which emotions are most confused and are they consistently confused with the same

concept/graphical representations?

Design

The studies investigate people's perception of the nine emotions chosen for the *Sorémo* system. It included one within-subject independent variable with nine levels representing the emotions to classify, and one dependent variable representing how each card was classified.

Participants

The redesigned pictorial representations of the emotions, was first tested using adults since the method is to be used by both teachers and pupils. The second and third iterations involve French, and then French and English children, respectively.

Procedure

Each participant was provided with a pile of randomized cards containing the nine cards from the *Sorémo* system. They were requested to peruse each card in turn and to place it on the emotion they thought it best represented on the *Sorémo* board in front of them, as illustrated in figure F.1.

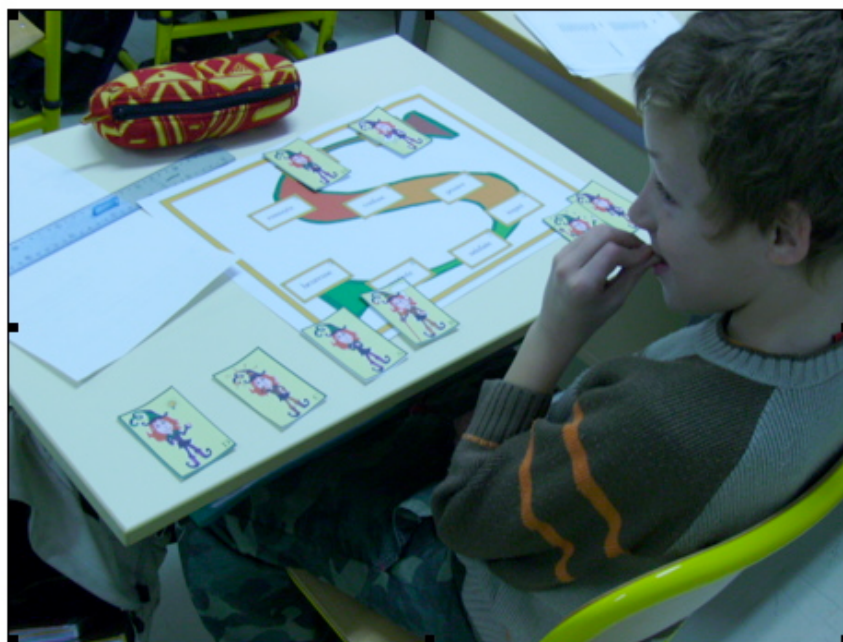


Figure F.1: The Design Iteration Procedure

They were allowed to take back a card already placed on the board to place it somewhere else if they changed their minds. Also, several cards could be placed on one emotion, and it was possible for emotions not to have any associated cards.

As illustrated in chapter 2, the *Sorémo* board consists of nine boxes on a S-shaped scale with the emotions written in the centre of the boxes. The scale uses the Traffic-light system, going from the most positive emotion, on the bottom-left hand-side in green, to red on the top-right hand-side for the most negative emotion.

ITERATIONS OF DESIGN

First set of emotions

In order to choose the set of nine emotions to represent in the new method, a set of 15 children (7 French, 8 English, aged 7 to 10 yrs) and 12 adults (international) were requested to draw a list of the emotions they would prefer to use to express their emotional states whilst playing educational games. The children were presented with a set of short educational systems to interact with, and were instructed to annotate the emotions they were feeling whilst playing. The adults were requested to consider the 14 different emotions present and to draw a list of any emotions they thought were missing to express emotional states whilst playing a game. All participants were then instructed to select four positive, four negative and one neutral emotion for the new method.

The four positive emotions chosen most often were *amusement*, *happiness*, *satisfaction* and *inspiration*. However, the first reported study showed that children have great difficulty in recognizing *satisfaction* and *inspiration* (with a percentage of correctly placed cards of 24% and 27% respectively), and they confuse them in many cases with other positive emotions. *Pleasant surprise* was therefore chosen for the third emotion, and the highest positive emotion considered to be missing from the set, *captivated* was chosen for of the fourth positive emotion.

The four negative emotions most requested were: *boredom*, *disappointment*, *unpleasant surprise* and *anger*. However, the children who were requested to verbalize their emotions when working on educational applications rarely used the labels *unpleasant* or *unpleasant surprise*, which was replaced most of the time by *puzzled*. Those two emotions were usually elicited when the child did not answer a question properly. In that case, they expressed puzzlement as to why they did not have a correct answer, as opposed to *unpleasantly surprised* by the result. For this reason, we chose to replace *unpleasant surprise* by *puzzled* in the set of nine emotions for use in the new method. Thus a first version of the Sorémo method was designed with a set of nine emotions presented in Table F.1.

Table F.1: the nine emotions of *Sorémo*, first iteration of design

Positive emotions	Neutral emotions	Negative emotions
Captivated	Thoughtful	Puzzled
Happy		Bored
Pleasant surprise		Disappointed
Amusement		Angry

The pictorial representations of the emotions were redesigned to correspond to the emotional expressions most recognized in the PrEmoTM in the first study. The character chosen to portray the emotions is a young witch, which will also be used as a theme for the fantasy game *DividingQuest*. The expressions were completed making full use of the Traffic-Light metaphor in the design of the witches' clothes, as illustrated in figure III.X.



Figure III.4: The evolution of colours in the witches' clothes.

Increasing their emotion representation capacity, the Traffic-Light metaphor was embedded in the design of the clothes to emphasize the valence of the emotional state represented: the more the emotion represented is positive, the more the witch's clothes turn to green (resp. red for negative, and orange for neutral).

The pictures were also inspired by English and French cultural reference for the inclusion of objects (the light bulb coming out of a person's head for *inspired*, the witch's wand broken in *anger*, etc...), comic literature (Foster, 1891), and the design of game characters, (Isbister, 2006).

First design iteration

The future method aims also to be used by teachers as a component of the child user-model in an OLM system, so a design choice was made to perform the first iteration on a set of 25 international adults working at the University of Bath.

In summary, the study results, show that negative and neutral emotions were well recognized, with a percentage of correctly placed cards ranging from 59% to 81.5%. The emotion *happy* was also well recognized (78%).

However, the other three positive emotions (*captivated*, *amusement*, *pleasant surprise*) were poorly recognized. This resulted in the pictorial representation of *captivated* being modified and *amusement* and *pleasant surprise* being replaced by *satisfaction* and *inspiration*, the next most accurately recognised emotions from the first study.

In order to enhance the recognition of those concepts, we redesigned the graphical representation of those three emotions in participatory-design sessions with English children.

Second design iteration

As the graphical representation of the emotions were designed in collaboration with English children, we performed the second design iteration of the Sorémo method with 27 French children, aged seven to nine years. This was to assess whether the graphical representations could be recognized by the French children, limiting recognition problems due to cultural differences.

The results of percentage of correctly placed cards during the study showed that whilst the emotions *thoughtful* and *puzzled* were confused for one another, the positive emotions *captivated*, *satisfied* and *inspired* were confused with the other positive.

Third design iteration

The positive and neutral emotions were redesigned with a group of both French and English children. This process was accomplished by adding some meaningful objects corresponding to each emotion, and more coloured objects following the Traffic-Light metaphor, in the graphical representation of cards. This iteration of design was then tested on a group of 10 English and 10 French children aged 6 to 12 years. The percentages of correctly placed cards ranged from 70.4% (*captivated*) to 100% (*angry*, *boredom*).

CONCLUSION

The iterative participatory-design process produced a set of 9 digital characters representations of emotional state, one for each of the following states: *happy*, *captivated*, *satisfied*, *inspired*, *thoughtful*, *puzzled*, *boredom*, *disappointed*, and *angry*. While emotions from the same polarity were easily confused with one another in the PrEmo system, they seem fairly well identified under this affective-and-coloured representation.

The facial/bodily expressions as well as gestures used to emphasize the emotional value of the digital representations might be more easily identifiable by children due to the participatory aspect of the design: the graphical representations were built from the comments

made on the PrEmo system representations in section 3.2, and involved children in the entire process in terms of usability testers, but also as full design partners. Initial sessions aimed at a better definition of the French and English common representations of the different emotions considered. Children were asked questions such as “When you think of a person feeling angry, what is s/he going? How does s/he look?”

The affective cues were embedded in all body parts for each emotional state, as well as the character’s clothes, and put in situation by the addition of objects the witch would be ‘familiar with’. The body proportions are not kept with regards to the head size, in order to enhance the facial cues visualization. However, unlike in the PrEmo system, the child testers were not put out by this infraction to the real world’s view of body stature. Indeed, the character’s hair and hat form somewhat a believable figure, and contrary to the child represented in the PrEmo system which brings the user closer to reality settings, this witch already belongs to the fantasy world where, by definition, anything can happen, even having a big head.

The use of objects to put the affective states in context was found really useful to help children’s emotional representation identifications, in particular with the lightning bulb coming out of one’s head for *inspiration*, the wand breaking in the witch’s hand for *anger*, or the question marks coming out of the witch’s hat when she finds herself *puzzled*. The items, described and designed by the child design partners, were taken from the comics and manga literature, quite popular for French children aged 6 to 12 years.

When concerned with simulating affect in pedagogical software, a main research question concerns the choice of emotions to simulate, and how to represent them graphically. In this PhD’s work, Embodied Pedagogical Agents will be used as interaction medium between the user and his/her learner model, by the means of affective feedback. The next section presents the design process that led to the definition of an emotional model of EPA’s affective responses to software interaction, for mathematical drill-and-practice applications, aimed at French and English children aged 6 to 12 years.

APPENDIX G: ASSESSING SYNTHETIC “INTERFACE CHARACTERS” EMOTIONAL STATE

AIM OF THE STUDY

This study aims to investigate the understanding of the instrument by children when asked to represent an interface character’s emotional state. Also of interest is to investigate whether the Traffic Light metaphor (using colours) aids understanding, and if there are any cultural differences between French and English children.

PARTICIPANTS

The first evaluation session was conducted with a class of 26 English children aged ten to eleven years. The second evaluation session was performed on a class of 29 French children aged eight to ten years. Due to the order in which the animations were shown, and the need to have matched pairs of animations (Traffic Light, colour and no Traffic Light, no colour) to analyze the data. The analysis could only be performed on complete folders of 10 pages. For this reason, one English and four French participants were discarded. The analysis was therefore run on two groups of 25 children.

SET UP OF THE EVALUATION SESSION

A video projector was utilised to display the animation in one single evaluation, with the children’s responses presented in a folder containing 10 pages (5 pages with, and 5 without, the Traffic Light metaphor). Each page contained a coloured picture of the Sorémo board and a label ‘Animation #’, on the top-right corner, numbered from one to ten. The animations were randomly mixed and presented to the children in the order illustrated in Table G.1.

Table G.1: The 10 graphical animations

Animation	Emotional states	Order	Animation
A	Disappointment	1	A Coloured
		2	B
		3	C Coloured
		4	D
		5	E Coloured
B	Unpleasant surprise	6	A
C	Inspiration, Amusement	7	B Coloured
D	Boredom	8	C
E	Satisfaction, Admiration	9	E
		10	D Coloured

In the animations with a letter only (A, B, etc...), the emotional state of the interface character (witch) was pictured with facial and bodily expressions and verbal cues. In the coloured animations (A Coloured, B Coloured, etc...), the Traffic-Light System metaphor was used to emphasize emotional state: a flow of colours through the witch's clothes according to the position in the emotional dimension 'pleasantness'. For example, in the animation 'C' representing the emotions of inspiration and amusement, a flow of green appears on the witch's clothes from bottom to top, (from her boots to her hat).

The animations were not based on sound or words, enabling us to conduct the studies in both France and England. The animations were replicated three times. The children had one minute to scale the emotions on the *Sorémo* grid, corresponding to the animation, by placing a cross on the point-scale that represented the strength of the emotions represented by the character.

In order to reduce the length of the experiment and therefore the cognitive effort asked from the children, only a sample of the emotional states were tested, all part of the emotional model presented in chapter 2. for the positive emotional states, two emotions are presented, one after the other: for example the witch is first *inspired* by the way the questions have been answered, and then *amused*. The research aim was to identify which emotional state would be considered portrayed by the witch after such animations.

RESEARCH QUESTIONS

The experiment aimed at answering the following research questions:

- Do children recognize correctly the emotional states represented by digital characters 2D animations? Does the presence of colours following the Traffic-Light metaphor help in the identification of the emotional states represented?
- Is there a cultural difference in the ability of children to recognize the emotional states of digitally represented characters?
- How do children use the instrument to represent emotional state of digital characters? Is there a cultural difference in the system's scaling procedure?

RESULTS

RECOGNITION OF EMOTIONS.

The results were first quantified according to the accuracy of recognizing the emotional states in the animation for each class (French and English). The two evaluation sessions were then combined (cf. Table G.2).

Table G.2: Classification of recognized emotional states by nationality and accuracy (percentage)

Emotional State	French		English		Total	
	%	SD	%	SD	%	SD
Disappointment	90.4	0.501	92.3	0.502	91.3	0.501
Unpleasant surprise	80.8	0.769	89.4	0.768	85.1	0.768
Inspiration, Amusement	76.0	0.707	85.6	0.790	80.8	0.748
Boredom	86.5	0.500	94.2	0.502	90.4	0.501
Satisfaction, Admiration	80.8	0.844	85.6	0.827	83.2	0.835

There is little difference in how well French and English children recognized the emotions, the English children recognizing them better with an improving factor ranging from 1.02 to 1.12. Chi-square tests performed on the total number of emotions recognized, and then considering each emotion separately failed to reveal any cultural difference between the French and English children (cf. *section 3.3.1* in the appendixes). Other tests comparing the recognition of all emotions, and then each emotion separately, also failed to reveal any significant difference in the age of the participants (4 groups were considered: 8 yrs, 9 yrs, 10 yrs, and 11 yrs old).

The emotional states were on average well recognized, with a percentage accuracy ranging from 76.0% to 94.2%. The positive emotional states were less well recognized than the neutral or negative ones, followed by the negative emotional state *unpleasant surprise*.

However, those three emotional states were a combination of several emotions, whereas *disappointment* and *boredom* could be associated with a unique emotional card on the method. The children confused emotional states from the same emotional polarity. For example, they confused *disappointment* with *puzzled*, instead of *happy* or *captivated* that belong to the other polarity of emotion. Only two children from the English group confused some emotional states with the opposite polarity, all on the animation *unpleasant surprise*.

IMPACT OF COLOURS.

Table G.3 represents the percentage of accuracy in recognizing the emotional states pictured in the animations. It separates the results according to the presence or the absence of colours following the Traffic-Light system metaphor in the clothes of the fictional user interface character.

Table G.3: Classification of recognized emotional states by (no)presence of colours (percentage)

Emotional State	Coloured		Not Coloured	
	%	S.D.	%	S.D.
Disappointment	49.5	0.37416	39.9	0.49453
Unpleasant surprise	38.9	0.78951	61.1	0.81649
Inspiration. Amusement	62.5	0.71181	37.5	0.63702
Boredom	83.9	0.40825	16.1	0.38069
Satisfaction. Admiration	54.4	0.64549	45.6	0.64689

A matched pairs Wilcoxon signed rank analysis revealed that children seemed to better recognize emotions when presented with ‘coloured’ animations than with the ‘non coloured’ ones ($z=-5.362$, $p=0.001$), with an improvement factor ranging from 1.2 for *disappointment* to 5.2 for *boredom*.

However, one emotional state *unpleasant surprise* is not as well recognized with colours as without, with a factor of 0.6. This emotional state resulted in a better score in the ‘partially accurate’ answers (corresponding to the middle scale, *orange*) with the colours than the non-colours. This may imply that this emotional state produced more confusion as to what the emotion was, leading the children to rate the emotion with less conviction and confidence.

McNemar tests investigating the impact of colour on the recognition of each emotional state revealed the difference to be significant for 4 out of the 5 emotional states: *disappointment* ($\chi^2=9.345$, $p=0.012$), *inspiration-amusement* ($\chi^2=11.429$, $p=0.001$), *boredom* ($\chi^2=31.688$, $p<0.001$), and *satisfaction-admiration* ($\chi^2=26.694$, $p<0.001$).

USER STRATEGIES IN USING THE SORÉMO INSTRUMENT.

During the study, five strength Strategies of use were identified when considering how children use the method to scale the emotional state of the characters. One example of each strategy is illustrated in Figure G.1. These strength Strategies of use can be identified for each child from the study data, and were quantified in Table G.4.

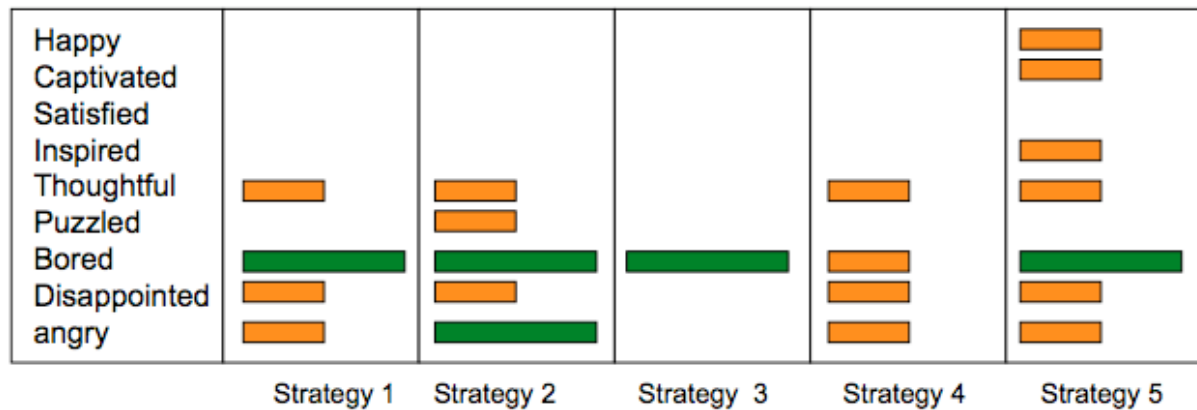


Figure G.1: Strength Strategies of use when filling in a *Sorémo* Screen

Strategy 1: Dominance and uncertainty - children following this behaviour chose one emotion with certainty (colour green) and then add a few emotions from the same polarity of emotions (uncertainty).

Strategy 2: Complex view of emotional state – children consider that the animations present several emotions at once. They put a green cross on two or more emotions, and then complete a few others in orange from the same polarity.

Strategy 3: Emotional state as a unique emotion – children consider the emotional state of the fictional character can be identified by only one emotion for each animation. They choose the emotion they are sure the character expresses and chose the scale ‘green’.

Strategy 4: Indecision over the emotion to choose - children are unsure which emotions are represented. They select a few emotions, usually from the same polarity as the emotion portrayed, and scale them as orange. This behaviour can be similar to the users of smiley-o-meters or quality charts that ‘go for the middle answer’ (Bull et al 2005).

Strategy 5: Wrong Polarity – children scale some emotions with the correct polarity as orange and green, and then rate some of the emotions of the opposing polarity in orange. For example, in Figure G.1, the child rated *boredom* in green, *disappointed* and *angry* in orange, and then *inspired*, *captivated* and *happy* in orange.

Table G.4: Classification of participants by Pattern (number of subjects)

Nationality	Nb	Pattern 1	Pattern 2	Pattern 3	Pattern 4	Pattern 5
French	25	6	8	9	2	0
English	25	9	3	5	6	2
Total	50	15	11	14	8	2

Children followed Pattern 1 the most when scaling emotional states (15/50 in total), closely followed by Strategy 3 (14/50). The ‘most-followed’ Pattern in the English group is Pattern 1 (9/25) was the ‘most-followed’ Strategy for the French group (9/25). Where the English children appear to be uncertain of the emotions (6/25 for Pattern 4), only 2/25 children seem to be indecisive about their answers. The French children seem to first conceive the emotional state as a unique emotion (9/25 for Strategy 3), quickly followed by the Pattern 2 (8/25) and 1 (6/25), considering the emotional state as a product of several emotions. In the English group, Pattern 5 (2/25) appears where children confused some emotional states with those of the other polarity.

DISCUSSION & CONCLUSION

The children were able to recognize, and express the emotional states proposed quite well using the *Sorémo* method. Little difference was found in the recognition of the emotional states. In general, the presence of the Traffic Light system or ‘colours’ in the digital character’s representation, reinforced the understanding of an emotional state.

To complete the investigations as to the potential of *Sorémo*, the next section presents the results of a pilot study on how children use the instrument to represent what they feel during a learning session on OLM pedagogical software.

However, when looking at the way children completed the *Sorémo* instrument to scale the emotions, some differences could be found between the English and French children populations: The English participants appear to represent the emotional states first and foremost as a product of complex emotions, with only 20% (5/25) of the children considering only one emotional state. The remainder (20/25) consider more than one other emotional state. On the contrary, the French group had a more equal share of complex and unique views. In the both groups, the youngest children were the ones following the unique view Strategy (8.5 to 9 years). This gives rise to the question as to whether the view of an emotional state as one emotion is more present in the age range of 7 to 9 years than for the older children.

APPENDIX H: ASSESSING YOUR OWN EMOTIONAL STATE, USING SORÉMO FOR THE FIRST TIME.

AIM OF THE STUDY

The research aim of this pilot study is to investigate how children use the *Sorémo* instrument when asked to represent how *they* feel: are they able to realistically represent their emotions, as they are experienced? Other questions to be addressed include: do such emotions vary when playing an entertaining educational game, and can it be portrayed through the instrument? Do they fully understand the method?

PARTICIPANTS

The study, performed in a French primary school, involved 18 children aged 7 to 11 years old and 3 HCI evaluators. The investigations on the instrument's usability raised the question of a developmental difference as opposed to a cultural difference in the use of the method. For this reason, only one culture, French children, was represented, with two age groups: 7-9 and 10-11 years olds, to investigate the developmental difference.

PROCEDURE

The children undertook the study individually, three at a time, in order for each evaluator to observe one student at a time. Each child was asked to sit at a computer and play a mathematical game for 15 minutes. The software chosen for this experiment, Multiplotest 1.0, is an educational game with an open-learner model that aims to teach children how to multiply small, and then large numbers, as well as practising their multiplication tables. All children were familiar with the game, having used it from January to June 2008 in a longitudinal study, described in chapter 4. The *Sorémo* method, however, was completely new to the children.

When accessing the computers, the children were presented with the *Sorémo* instrument, and asked to express how they were feeling at that moment (considered as the before test). They were then free to use all features of the game at their convenience. At various moments of the interaction (from a list of learning situations to investigate, when a strong child emotional display was detected), the evaluators prompted the children to click on the '*Sorémo* button', and give their emotional state by using the instrument integrated to the game, taking note of what lead to this/these momentary emotion(s). The children were also free to use the

Sorémo interface whenever they wanted during the session. At the end of the session, a final *Sorémo* screen appeared (after test), which the children had to complete. They were then asked to complete a small questionnaire on the usage and usability of the instrument.

MEASURES

The *Sorémo* instrument provides researchers with a series of quantitative measures taken from the user's interaction with the interface:

- o TFirst, the number of seconds spent until the user scaled the first emotion. This represents the time the children took to take in the software features and think about their answer. This measure should show whether there is an exploration time during the first use of the software, or if children manipulate the interface without any processing time.
- o TLast, the total number of seconds taken to scale the nine emotions. This qualifies as the actual time of interaction with the nine emotions, thinking about, and then scaling the emotions. From this measure, a third measure was calculated, TInteract, representing the amount of seconds spent actually scaling the emotions, which corresponds to the formula: TLast-TFirst.
- o and for each emotion the history of scaling. This shows the number of changes made while scaling, as well as how those changes were made.

Each child was given a questionnaire at the end of the session asking them to rate by means of a 5-points Smiley-o-meter scale the extent to which: they understood the emotions displayed, whether the set of emotions represented what they felt while playing, how helpful the use of both metaphors were for understanding the meaning and valence of the emotions, whether they could correctly perceive what scale they chose, and their willingness to use such an instrument in an educational game.

RESEARCH HYPOTHESIS

The following research hypothesis were investigated:

H₀¹: There is no difference between the number of seconds spent before first interacting with the instrument (TFirst) between the three moments of interaction.

H₁: Children spend more time before first scaling the instrument when using the instrument for the first time than in the 'during' and 'after' instrument completion.

H₀²: There is no difference between the number of seconds spent scaling the instrument (TInteract) between the three moments of interaction.

H₂: Children spend more time scaling the instrument when using the instrument for the first time than in the ‘during’ and ‘after’ instrument completion.

H₀³: There is no difference between the number of changed registered during software use between the three moments of interaction.

H₃: Children made more changes in the scaling of the emotions when using the instrument for the first time than in the ‘during’ and ‘after’ instrument completion.

At last, the following research questions were considered:

- Is there a developmental difference in the use of the instrument to scale one’s emotional state?
- How well are the digital representations of emotional display understood? How well do the pictures represent what children can feel during a pedagogical session on OLM software?

RESULTS

COMPLETION TIME OF THE SORÉMO TEST

Two measures registered by the instrument include the time of response for children to complete a test: *TFirst*, the number of seconds before the first emotion is scaled; and *TLast*, the total number of seconds spent to scale the nine emotions. From this data, a third measure was calculated, *TInteract*, which represents the actual time children needed to scale the nine emotions after scaling a first emotion. These measures were taken during the before test, the intermediate ones, and the after test. The average time of response each test is represented in Table H.1:

Table H.1: Time of first response and total response (number of seconds)

Test	TFirst		TInteract	
	Mean	S.D.	Mean	S.D.
Before	25.72	14.438	30.222	13.871
During	10.61	5.089	21.389	11.683
After	7.17	4.162	18.278	10.191

Overall, children spent more time completing the procedure during the first interaction with the method (before test), than while playing Multipliotest (average of 48% seconds less), and at the end of the session (average of 78.5% seconds less). The time for the first scaling (TFirst) is also reduced considerably (average 66%).

First Scaling Time

The results of a Friedman's chi-square test show a significantly reliable difference in the number of seconds spent before first interacting with the instrument (TFirst) between the three moments of interaction ($\chi^2=23.324$, $df=2$, $p<0.001$), rejecting H_0^1 . Children spent an average exploration time (in seconds) going from 25.72 ± 3.403 as "Before session" interaction, to 10.61 ± 1.199 "during the software use" interaction, and 7.17 ± 0.981 as "After session" interaction.

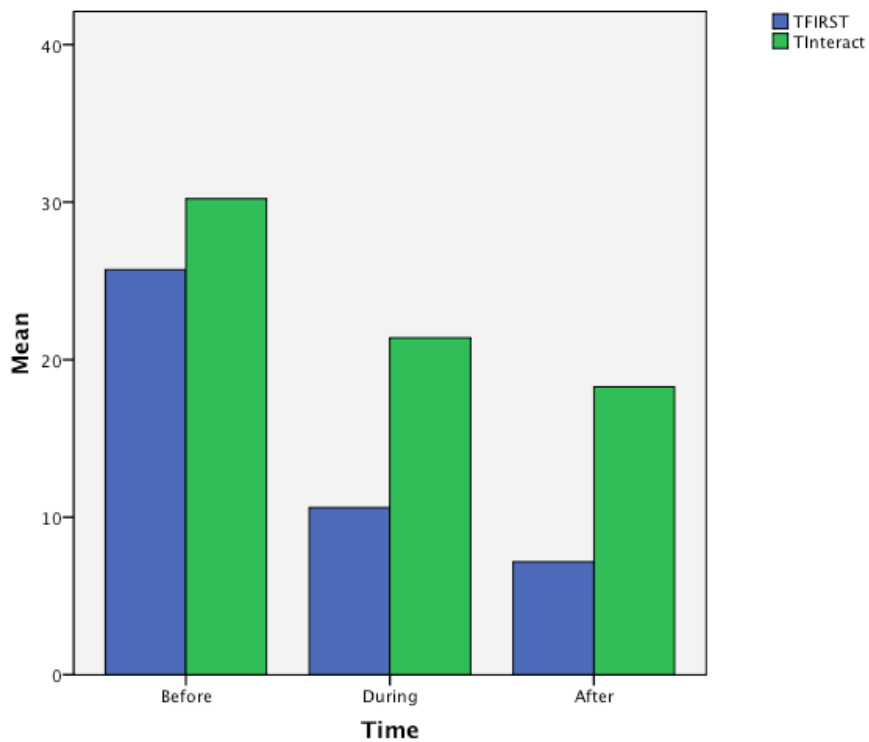


Figure H.1: Average of First interaction and Total interaction time with Sorémo

This may support the evaluators' observation of a change in the children's scaling process: In the first test, children first study the instrument. They observe how it works, try to understand what the pictures represent, and think about their feelings before scaling the first emotional representation of their choosing, which brings TFirst to a total of 9 to 56 seconds (average 25.72). They then scale the emotions going from one picture to another and sometimes back, thinking about all emotions before beginning to answer. In the following tests however, they seem to interpret the emotional display more quickly, and scale the first emotion faster (66% in average), going from 2 to 15 seconds (average 7.33). While individual differences in answer time could be observed for the first report completion ($SD_{\text{before}}=14.436$), no participants were found to spend more time in the *during* and *after* tests than during this

first interaction with the instrument. This suggests practice effect that is unrelated to design of the software.

Scaling Interaction time

The results of a Friedman's chi-square test show a significantly reliable difference in the number of seconds spent interacting with the instrument (TInteract) between the three moments of interaction ($\chi^2=8.343$, $df=2$, $p=0.015$), rejecting H_0^2 . Children spent an average interaction time (in seconds) going from 30.222 ± 3.270 as "Before session" interaction, to 21.389 ± 2.754 "during the software use" interaction, and 18.278 ± 2.402 as "After session" interaction.

The significant difference may be explained by the students' interaction logs of another change in the children's scaling process: At first, they seem to consider the presence of all 9 emotions at once, before selecting the appropriate scale emotion by emotion. In the intermediate and after tests, they seem to think about the presence of each emotion iteratively, thinking about the presence of one emotion only once they scaled the previous one. They then look at the method in its whole, and adjust the scales of specific emotions when needs arises. This may suggest a recognition effect in the scaling process, with children recognizing faster the different emotions to be scaled, which would facilitate their reflection process as to the scaling completion.

CERTAINTY OF ANSWERS

The results of a Friedman's chi-square test show a significantly reliable difference in the number of changes made by the children when scaling the instrument (NbChange) between the three moments of interaction ($\chi^2=8.000$, $df=2$, $p=0.018$), rejecting H_0^3 . Children made an average of changes going from 2.50 ± 0.706 as "Before session" interaction, to 2.11 ± 0.471 "during the software use" interaction, and 0.89 ± 0.254 as "After session" interaction.

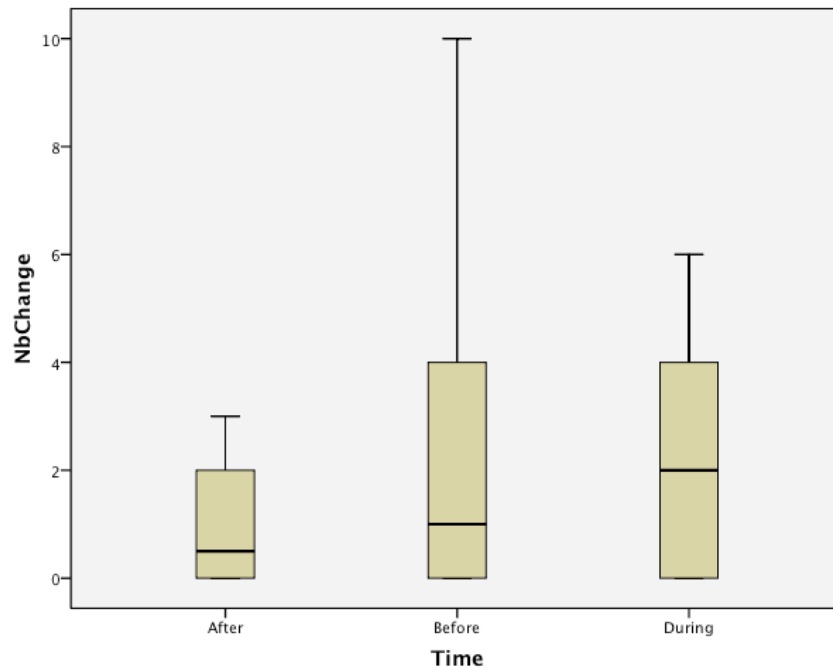


Figure H.2: Number of Changes according to the time of interaction

Children appear to be sure of their scaling of the pictorial representations of emotions: some children had no problems scaling the emotions in a first go (6/18), or only made a few changes on a couple of them (9/18), with an average of number of changes per emotion per test of 2.17. The two children who truly had some trouble in recognizing emotions, and expressed it verbally to their evaluator during the session (7 and 8 pictures out of 9), used very little number of changes when completing the intermediates and after tests, bringing down the number of changes per test to only a couple of changes for 1 to 3 emotions per test. Results from the evaluators' observations showed that only two children (2/18) chose to answer the pictures they were unsure about at random, while the others took more time to analyze the picture and gave an answer only once they reached a decision upon their meaning or strength.

USER COMPLETION PATTERNS OF THE USE OF SORÉMO.

During the study, four strength completion strategies were identified when considering how children use the instrument to scale their emotional state, which correspond to the four first patterns of use described in the previous study (*appendix G*). These strength patterns can be identified for each child from the study data for each use of the method, and were quantified in Table H.2, and then graphically illustrated in Figure H.3:

Table H.2: Classification of participants by Strategy (number of subjects)					
Test	Total	Pattern 1	Pattern 2	Pattern 3	Pattern 4
7-9	9	1	0	3	5
10-11	9	2	5	0	2
Total	18	3	5	3	7

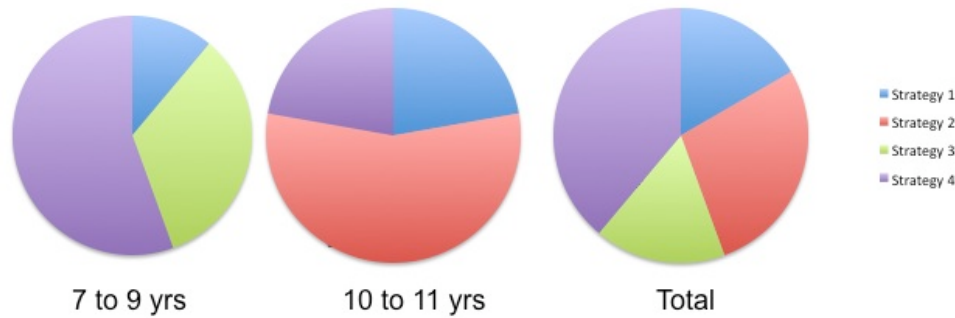


Figure H.3: Repartition of Strength Pattern of Instrument use

While the younger children have a tendency to being uncertain of their feelings (Pattern 4, 5/9), the elder group tends, while still viewing emotions as complex, to rate a couple of emotions as certain, and then a few in orange (Pattern 2, 5/9). It is also to be noted that only younger children viewed their emotional state as a unique emotion (Pattern 3, 3/9), the others considering emotions as a complex set of basic emotions (Pattern 1, 2, 4: 15/18). The notes from the observations of evaluators revealed students following Pattern 4 to be cautious in every action in the game, doing everything in moderation and with insecurity. After the game, they seemed more secure in their choices, and produced a set of more definite answers, rejoining the Pattern 2.

A various range of emotions were perceived during the game and represented with the method as a combination of existing emotional representations, and verbally expressed by children to the evaluators as: epiphany (10/54), confidence (21/54), confusion (5/54), frustration (15/54), or dissatisfaction (3/54). For example, the tests showing frustration were identified by the evaluators as prompted directly by the children experiencing difficulties with the point-select-and-click interaction style (use of a sliding bar to give an answer). They told the evaluators they wanted to *tell Multiplio* (the interface EPA of the game) about the problem they were experiencing and how frustrating it was to know the answer but not be able to give it right.

USABILITY OF SORÉMO.

Table H.3 illustrates the answers given by the children on the usability of the method after the experiment, by means of a 5-points smiley-o-meter: children rated each question, by selecting a smiley, with a scale illustrated here from 1 (does not agree) to 5 (totally agrees), and verbally justified their choices to the evaluators.

Table H.3: Answer of the smiley-o-meter (nb answer)

Test	Rating 1	Rating 2	Rating 3	Rating 4	Rating 5
Question 1	0	1	0	7	10
Question 2	0	1	2	1	14
Question 3	0	2	0	2	14
Question 4	0	0	0	6	12
Question 5	2	0	0	2	14

- Question 1: *How well do they understand the emotional representations in Sorémo?*

The pictures in *Sorémo* seemed well understood by the children with only one child rating them as quite difficult to understand. The children were asked to look at all 9 pictures and to tell the evaluator what they represented before scaling this question with the smiley-o-meter. A majority of them (10/18) claimed to have no problems understanding them, when the remaining seven struggled only on the meaning of a couple of pictures (*captivated* and *puzzled*), which correspond to the emotions the least recognized in the usability tests performed on the *Sorémo* method in *appendix G*.

However, a majority of the children (11 participants: 6 7-9 yrs old, and 5 10-11 yrs old) told the evaluators that their understanding of the emotional representations grew with time from the before to the after test; after reflecting several times on their meaning, they reached an understanding of what the pictures represented. Two trends could be identified, which were looked into for the TInteract measure: On the one hand, part of the children look through all cards and choose 1 or 2, considering each separately to check its correspondence to their current emotional state. On the other hand, the recall of the emotions included in the method makes them they decide in advance what their current state is, and look for the card that corresponds.

- Question 2: *How well do the pictures represent what children feel while playing at Multipliotest 1.0?*

Children mostly found the set of emotional representations to be comprehensive; as they could freely express the various emotions they were feeling while playing. Only one child identified one emotion, ‘insecure’, to be missing from the set.

- *Question 3: Did the use of colours in the instrument's background and witches' clothes help understand the pictures?*

A majority of children (16/18) found the Traffic-Light System metaphor helpful in identifying the emotional representations. Only two children were more reserved as to its benefits, as they did not notice the metaphor explicitly within the design of the interface.

- *Question 4: How well did the users see their answers when scaling?*

Most children stated that the scaling system was very explicit, easy to use, and showed appropriately what they had chosen as strength. 6 out of the 18 children were less enthusiastic than the others, wishing for the system to be even more interactive, with flashes of colour in the characters and the use of sound: While they were able to correctly explain how the scaling system worked to the evaluators, they wished for the background of the witches to stay of the colours scaled, or at least be more interactive to support the instant recognition of their scaling choice.

- *Question 5: Would they like to use the Sorémo instrument within mathematical software in the future?*

Only two children were not keen on using *Sorémo* as a self-reporting instrument. The others (16/18) expressed their willingness to tell the software how they feel, creating a sort of emotional dialogue with the software.

DISCUSSION AND CONCLUSION

Children seemed to take well to the instrument, with a majority of them willing to use the method in pedagogical software to input their emotional states at various moment of interaction. Most of the children associating with software's EPA *Multiplio*, stated their wish to be able to 'tell Multiplio' how they felt the next time they played the game.

In general, the scaling system was found easy to understand and use, and the children made effective use of all three points of the scale, using the middle scale to moderate their answers rather than avoiding to form an opinion.

There seem to be an exploratory and practice effect in using the instrument between the first and last *Sorémo* procedures. However, children seemed to feel confident they can use the instrument without any external help during the experiment, and that they 'knew how to do it' at the end of the session. Indeed, at the end of the experiment, they spent less time filling the 9 scales, with a time of first interaction with the instrument reduced to an average of 7 seconds, and a completion time of 18 to 20 seconds. This completion time seems in line

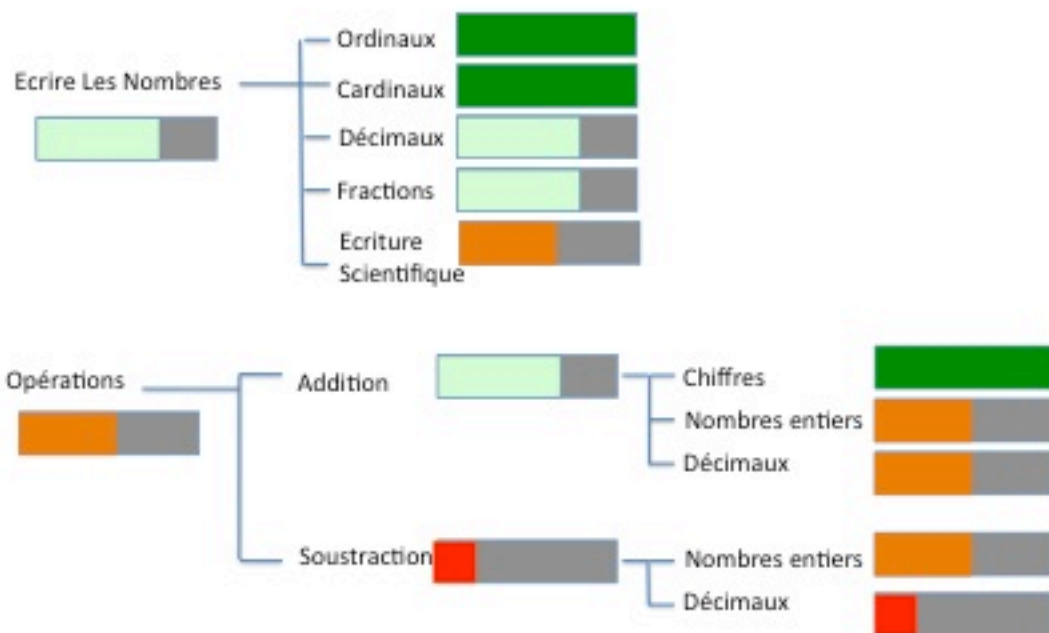
with the use of a self-report instrument, ie: it should not require too long a time for children to complete so as to distract them from the learning activity at hand.

Children feel more confident in their understanding of the digital representation of emotions after using the instrument a couple of times, where they could reflect on the meaning of the pictures through time to reach an understanding on what they represented. Children considered the set of 9 emotions to represent well the various emotional states they found themselves in during the learning session, and combined them to represent more complex states, by using the strength scaling to nuance the different emotions felt.

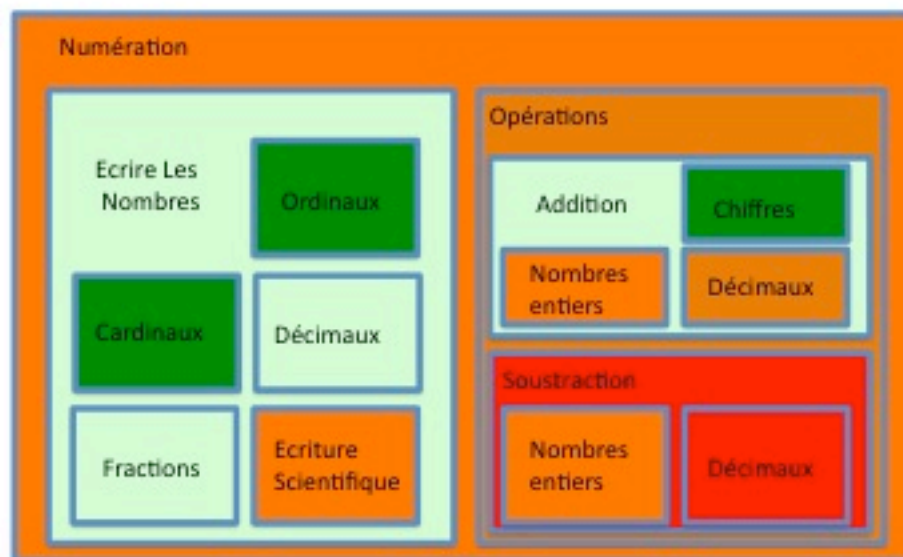
There seem to be a developmental difference in how children view affective states: From the strategies employed by the participants to scale their feelings, it can be seen that some children aged 7 to 9 still view any emotional state as a unique emotion, which corresponds to Izard's observations on young children's representation of emotional concepts for 5 to 8 years olds (Izard, 1984). On the contrary, all participants aged 10 to 11 involved in this experiment only viewed the emotional states as a complex set of basic emotions. This result seems to confirm the observations made on the *Sorémo* instrument completion for usability studies presented in the appendix G.

Appendix I: OLM representations

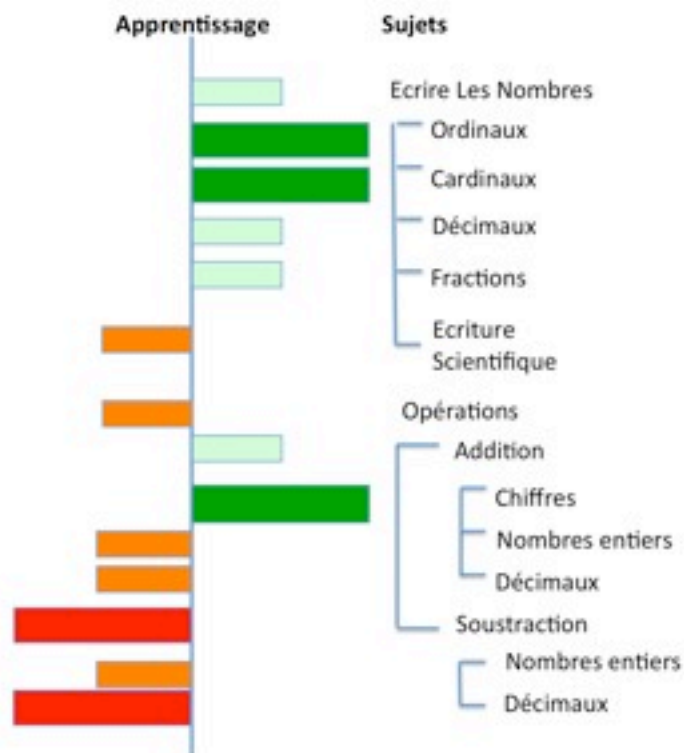
I.1 6 REPRESENTATION TOOLS UNDER THE “LECTURE VIEW”



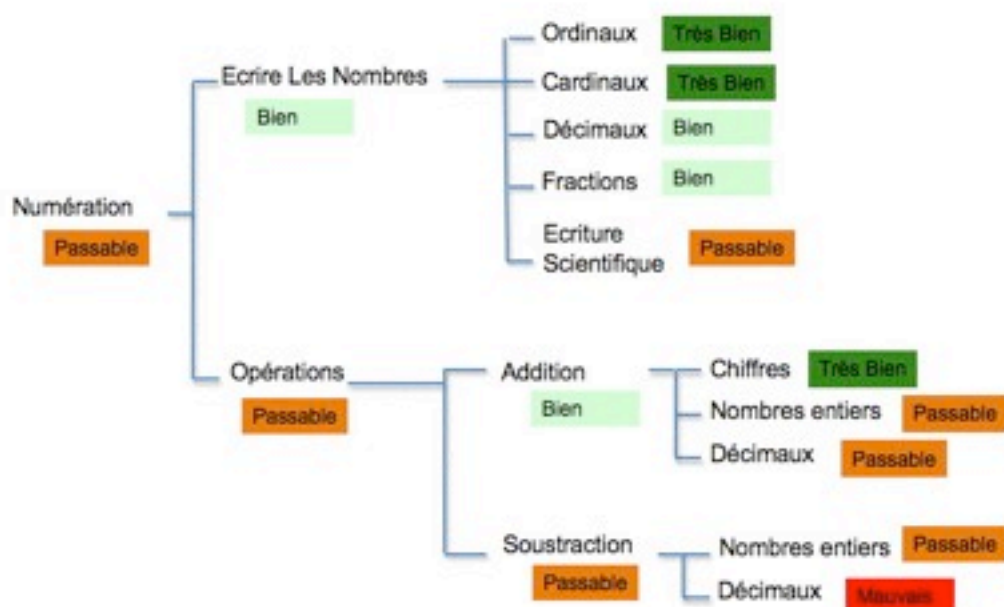
Skill-Meter



Boxes



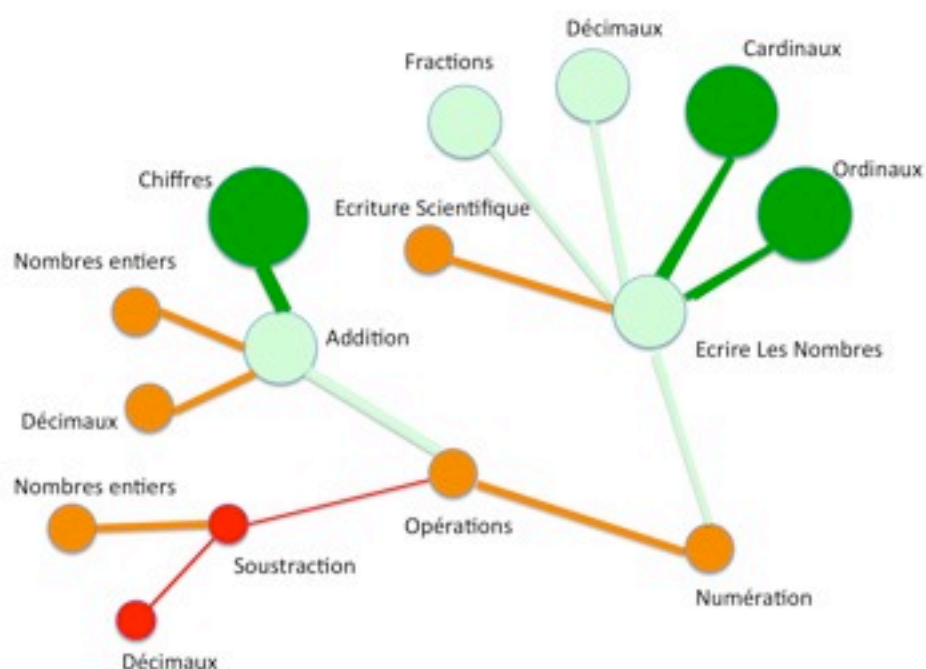
2D-Graph



Texte

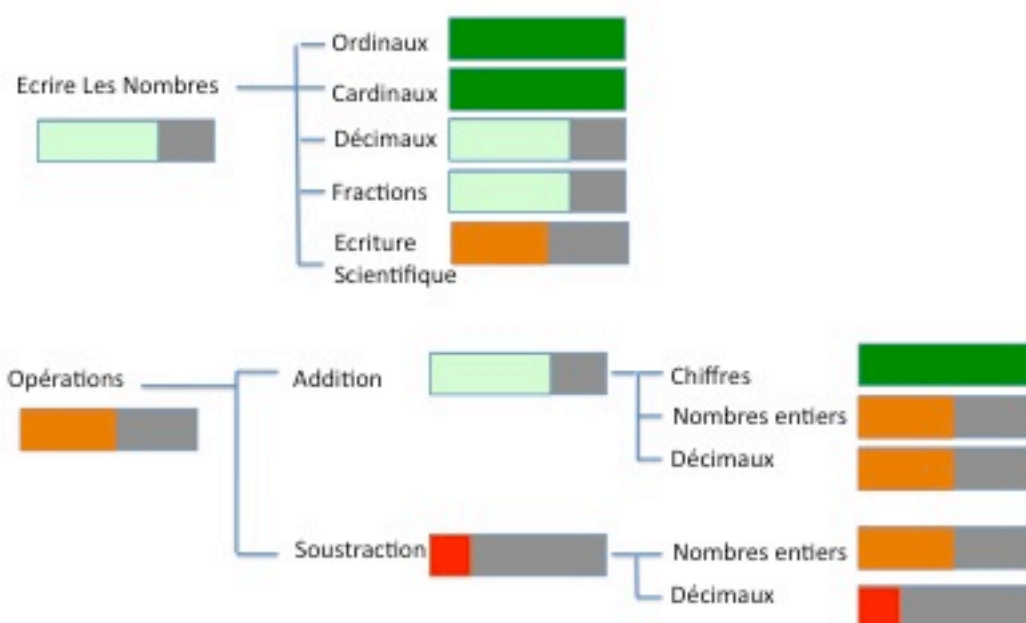
Niveau	Sujets à ce niveau
Très Bien	Cardinaux, Ordinaux, Chiffres
Bien	Ecrire les nombres, Fractions, Décimaux, Additions
Passable	Ecriture Scientifique, Operations, Nombres entiers, Décimaux, Nombres entiers
Mauvais	Soustraction, Division

Table

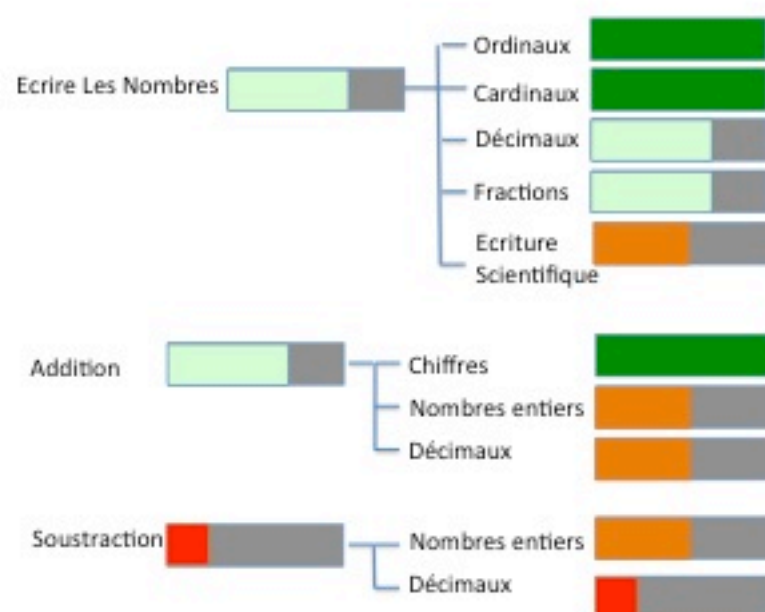


Bayezian

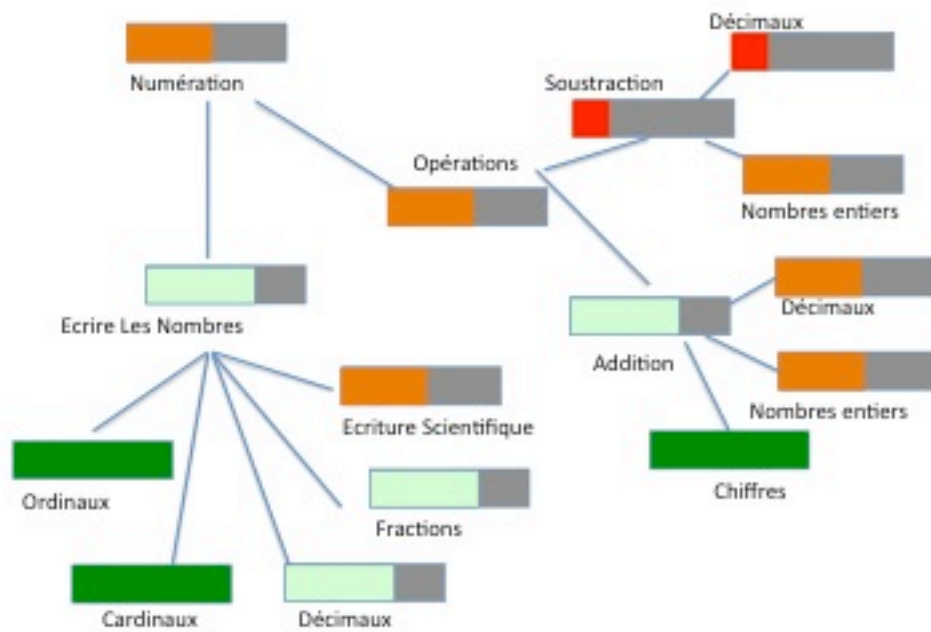
I.2 4 LM VIEWS FOR THE REPRESENTATION TOOL “SKILL-METER”



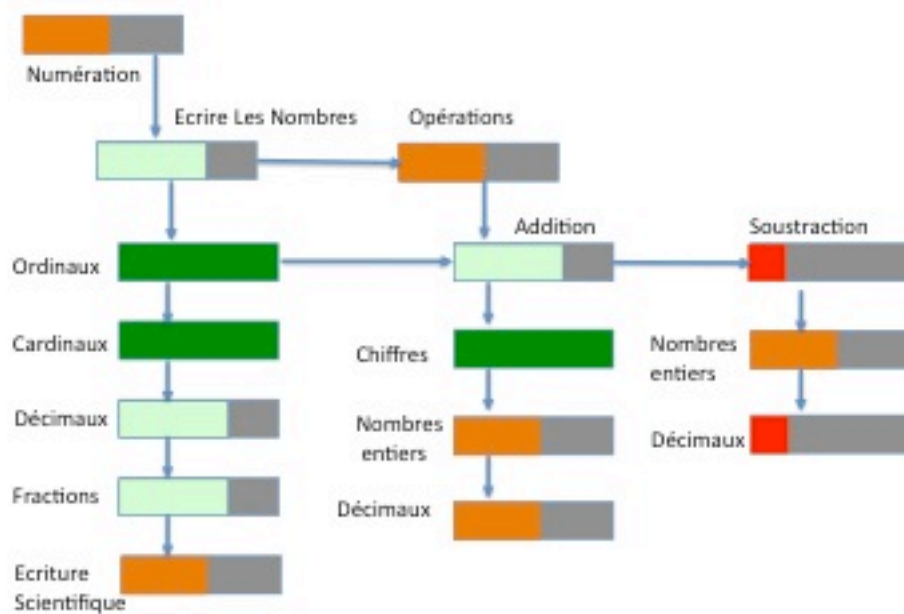
Lecture View



Related Concepts View



Concept Map View



Pre-requisites View

APPENDIX J: SMILEY FACES, TRAFFIC-LIGHTS, OR BOTH IN THE USER MODEL

AIMS OF THE EVALUATION

This study will investigate the impact of the inclusion of a coloured-metaphor of learning in the ability for children to visualize their knowledge acquisition, and ask for help at appropriate times through the visualization of their UM, and the interaction with the EPAs. Two sets of research issues are looked into, separated by research areas: affective EPAs, and use/perception of the user model.

COLOURED VS. EMOTIVE INTERFACE PERSONAS:

Aim: Investigating whether there is a difference in learning achievement and motivation between helping characters designed showing facial expressions, or helping characters designed showing facial expressions and using the Traffic-Light system.

Research questions:

- Do children perceive and appreciate the coloured metaphor of learning used in the design of the interface personas?
- Does the use of the traffic-light system as a learning metaphor encourage the child to be more motivated to learn ?

CHILDREN'S INTERACTION WITH AND PERCEPTION OF THEIR UM:

Aim: Investigating whether there is a difference in the way children perceive and interact with their LM of representing the model using the Traffic-Light system, an interface persona expressing emotions, or an interface persona expressing emotions and using the Traffic-Light system.

Research questions:

- Do children interact more with their UM in the coloured condition?
- Do children completely understand the possible use of OLM?

STUDY DESIGN

RESEARCH HYPOTHESIS

This section presents a list of the research hypothesis analysed in this study.

Learning occurring during the session:

H_0^1 : There is no difference in prior knowledge between the children in the control or coloured conditions.

H_1 : There is a difference between the pre-test scores registered for the children in the control or coloured conditions.

H_0^2 : There is no difference in learning gains between the open learner models that students interact with, when represented by the Smiley Faces metaphors alone, or with the two metaphors.

H_2 : The children learned more during the learning session in the coloured condition than in the other one.

Motivation:

For the study of the level of child's motivation, the following research question was addressed using graphs and tables as illustration of data tendencies:

- Is there a difference in children's motivation to use pedagogical software between the two conditions?

PARTICIPANTS

The experiment involved twenty-eight children of level "CE2-CM1" (fourteen aged seven to nine, and fourteen aged nine to eleven years old) and their teacher from a French primary school.

EXPERIMENTAL DESIGN

The experiment has been designed as a between-subjects comparative study with pre and post session learning tests. It includes a control and an experimental condition, differing

in the visualization of the learner model: The control condition is an open learner model where an EPA designed with the Smiley Faces metaphor of learning interacts with the user on the model inspection and edition. The experimental condition has an open learner model with an EPA designed using both the Smiley Faces and the Traffic-Light metaphors of learning.

EXPERIMENTAL MEASURES

The experimental study is designed as a between-subject comparative study across two conditions (each condition represented by a separate user-interface, one for each type of EPA). It includes both quantitative and qualitative measures registered during software interaction, after completion of pre- and post-session tests/questionnaires, concerning the different research questions to be answered on learning achievements, motivation to learn, and LM component.

LEARNING GAIN

An analysis to assess whether there is a learning improvement in each condition is realized by comparing children's pre- and post-session mathematical test results. This measure takes into account the improvement on learning on areas of the game, game activities that children may or may not have accessed during the learning session. To minimize threats to the test's validity, child-users were encouraged to avoid guessing in general, but especially in the pre- and post- mathematical tests, if they did not know or remember the answer to a question.

MOTIVATION

A second set of measures is taken concerning student's motivation in using the pedagogical system for learning, and their level of engagement within learning activities during the learning session. Qualitative measures are also taken on children's motivation to use the different systems, in post-session qualitative questionnaires using Likert scales.

PROCEDURE

The two interfaces of *Multiplotest* version 1.5 were used for this study, and differ only on the learner model features represented in navigation or mathematical help, and learner model visualization/access. Each child worked on a unique software interface, for a single session of thirty minutes.

The children worked in a classroom separate from their peers, in groups of three, in the presence of a researcher. The material was conveniently placed in order for children not to

see the others work, and all worked on the same experimental condition. They were therefore not aware of a difference in software interface between the children participating to the experiment until the end of the session, after completing the qualitative questionnaires.

After answering the computerized pre-test mathematical software, they were free to use software features to their liking for the rest of the allocated time. After answering the post-session computerized mathematical test, they were presented with a short questionnaire on their motivation to use software, and on their use of the OLM component.

CHOICE OF STATISTICAL ANALYSIS FOR RESEARCH HYPOTHESIS

Considering the number of research hypotheses to be statistically analysed and factors of the experiment, the level of statistical significance for rejection of the null hypotheses has been defined at $\alpha = 0.01$. Two research hypothesis, H1 and H2, were analysed statistically:

- The research hypothesis H1 aims at investigating whether the distribution of students into the control and experimental conditions resulted in groups of equivalent mathematical abilities when answering the pre-session test. A chi-square “goodness to fit” statistical test was performed on the children’s pre-session test scores, investigating the variance of the populations in their answers.
- The data gathered in order to investigate hypothesis H2 did not satisfy the assumption of normally distributed data, which is a requirement for any parametrical test such as ANOVA. Therefore, the analysis was performed using a non-parametric test, the Mann-Whitney two-sample rank-sum test, which tests for equality in medians.

RESULTS

Learning occurring during the session:

Figure V.1 outlines the results of the pre- and post-scores registered for each condition.

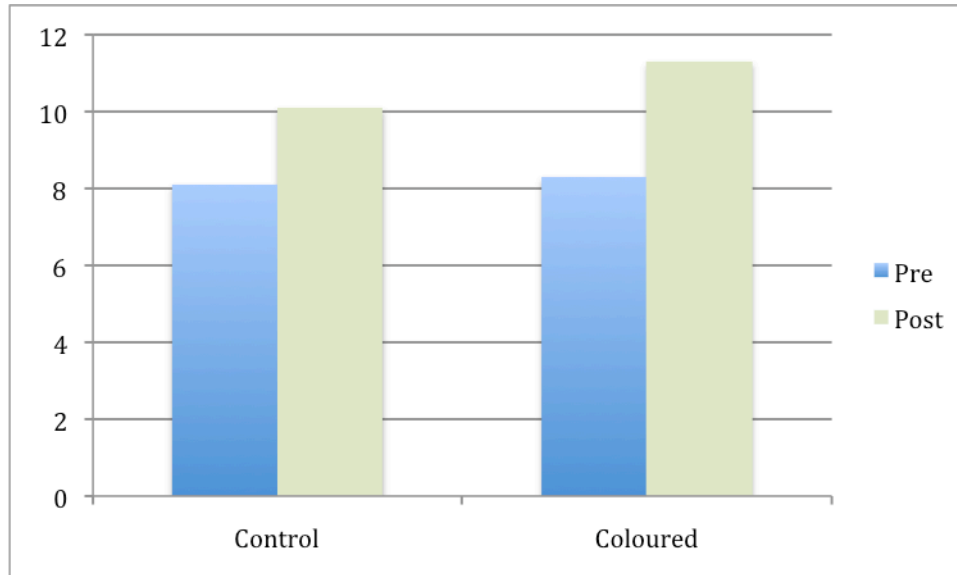


Figure V.1: pre- and post- mathematical test scores across conditions

A Chi-square ‘goodness to fit’ test failed to demonstrate a statistically reliable difference ($\chi^2=16.67$, $df=17$, $p = 0.719$) between the expected and observed frequencies in pre-test scores between the two conditions. The participants in both conditions seem to come from the same population, with a lack of statistical difference in prior knowledge across conditions. This means that inherent differences in the population would not be able to account for any differences found in the post-test, or the proportional learning gain within the session.






The results of the proportional learning gain (between the pre- and post- test scores) concerning the two conditions were compared using the Mann–Whitney two-sample rank-sum test.

Median latencies for proportional learning gains in the control group and the coloured-condition were 0.015 and 0.089; there is a statistical difference between the learning gains observed under the two conditions (Mann–Whitney $U=35$, $n1=n2=14$, $p=0.0013$ two-tailed). It can be concluded that the coloured-condition enables children to learn significantly more in terms of multiplication skills than the control one.

Motivation:

Table J.1 illustrates the answers given by the children on their motivation to use the system and level of engagement after the experiment, by means of a 5-points scale smiley-o-meter: children rated each statement, and selected a smiley that corresponded to their belief, with a scale illustrated here “does not agree” to “totally agrees”.

Table V.1: Answers concerning the motivation (nb answer)

					
Statement 1: Multipliotest is fun to use					
Control	2	7	2	3	0
TL	0	1	3	2	8
Total	2	8	5	5	8
Statement 2: With the game, I learned more about multiplications.					
Control	2	6	2	3	0
TL	0	2	0	2	10
Total	2	8	2	5	10
Statement 3: I would like to play Multipliotest again					
Control	2	9	3	0	0
TL	0	0	3	6	5
Total	2	9	6	6	5

There appears to be a difference in children’s rating tendencies concerning the level of enjoyment produced by the two different conditions: children using the coloured-EPA version seemed more satisfied with the application, experiencing more enjoyment out of the game, and visualizing more how it could be used to learn maths than the ones working on the control conditoi. A majority (17/28) of children found both software interface useful to learn multiplications.

Only one child did not consider the application enjoyable in the OLM condition, and did not find any use of the application for learning multiplications. She happened to be the youngest low-skilled student of the group. When interviewed, she expressed her frustration at playing this game, by not being able to remember her tables.

DISCUSSION

The study investigated the benefits of using an colours in the design of EPAs to help the understanding of the learner model component it interacts with.

While children taking part in the experiment were drawn from the same population between conditions, they performed differently while using the software: the ones working on

the coloured interface seemed to perform significantly better than the ones working with the control condition. Low-skilled children seem to benefit more from using *Multipliotest* in terms of learning gain than the other students, especially in the coloured-EPA interface.

However, such results should be investigated further, with children using the application in a longitudinal study to separate the impact of first time interaction with the game from children's behaviour on longitudinal software use. Indeed, in this experiment, children used software for a period of thirty minutes in both conditions, one time only.

Children working on the coloured version of *Multipliotest* seemed more motivated to use the application again in the future than the ones working on the control interface. They felt more concerned with what multiplication tables they knew, and how to “show the software” about their knowledge (current state, or newly acquired), and enjoyed playing with the game more.